The Science of Mind (精神科学) 52 (2014), pp. 243-260.

Language Death and Diversity: Philosophical and Linguistic Implications

Lajos L. Brons
Nihon University

In the first paragraph of "the Method of Truth in Metaphysics", Donald Davidson writes:

> In sharing a language, in whatever sense this is required for communication, we share a picture of the world that must, in its large features, be true. It follows that in making manifest the large features of our language, we make manifest the large features of reality. One way of pursuing metaphysics is therefore to study the general structure of our language. (1977, 199)

Much of this reflects (or even summarizes) Davidson's philosophical program, but much of it also reflects a (perhaps, defining) tendency in analytic philosophy as a whole: language, especially ordinary language, is to be taken as a guide to how things are.

For much of the 20th century, analytic philosophers have been referring to examples taken from the English language to prove their points, oddly ignoring the thousands of other living languages under the apparent assumption that all there is to be learned from language can be learned from just that one. In the last few decades this has been changing (a bit), and (a few) other languages are now (sometimes) taken into account as well. Linguists have been aware of the great variety among languages for much longer, of course, and although the Chomskyan research program of *Universal Grammar* was originally based on similarities between just a hand-full of languages, it is gradually extended and refined by bringing more and more languages into its scope. (On the other

hand, linguists starting from the other end – taking actual linguistic diversity as their starting point – often consider this Chomskyan project unsuccessful, pointing out exceptions for every suggested universality. See, for example, Evans & Levinson 2009.)

Although there can be little doubt that taking full linguistic variety into account is necessary if one wants to learn from language – either about language itself or about its philosophical implications – there are good reasons to doubt that this is sufficient. The question that needs to be asked is: How much can we learn from *actual* linguistic diversity? An approach that, for example, bases linguistic theories about possible and impossible languages or philosophical theories about existence and reality on actual linguistic diversity implicitly assumes that actual diversity more or less coincides with possible or total (historical) diversity. However, the set of currently living languages is only a subset of all languages that have existed throughout human history, and the smaller that subset, the more implausible the assumption, especially if there are reasons to believe that the subset is not a random sample.

For example, if it would be found that all living languages represent time (some by means of tense marking, some by means of adverbs, some otherwise), then one may be inclined to take this to be a universal feature of language, and perhaps, to consider this philosophically significant (for the metaphysics or phenomenology of time, for example, or for both). However, if there would be a dead language (*i.e.* a language that was once spoken, but that died out some time in the past) that in no way represented time, then this would refute such conclusions. This, of course, raises another question: How many dead languages are there, and what can we know about them? But that is really the same question as the one above.

The questions that need answers then, are: **(1)** How many languages have been spoken (or have existed) on earth? and **(2)** Are there (good) reasons to believe that the subset consisting of currently living languages is a representative sample (in the relevant respects) of the set of all languages that have existed?

There have been a few attempts to answer **(1)**. According to David Crystal (2002), for example, a reasonable estimate would be between 64,000 and 140,000 languages; and Mark Pagel (2000) calculates it by means of a complex formula, but gets different numbers depending on various assumptions: half a million if it is assumed that humans started speaking 200.000 years ago, or 130.000 if language began 100.000 years ago (and if language change was slower).

This paper tries to answer **(1)** by means of a different approach: modeling instead of calculation or estimation.[1] The advantage of this approach is that it simultaneously provides hints (at least) towards an answer of **(2)**. Results of this approach suggest that the total number of historical (dead plus living) languages is probably somewhere around 150.000 (plus or minus 50,000), but possibly much higher, and that the 5% or less of those that are spoken today are almost certainly *not* a representative sample thereof. Therefore, no universal claims about language and/or its philosophical implications can be based on *just* actual linguistic diversity. Nevertheless, even if actual linguistic diversity cannot tell us what is *maximally* possible (or necessarily universal), it can tell us what is *minimally* possible.

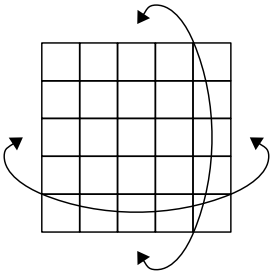
the model

Imagine a world populated by 25 cultural communities. In the model these are represented as 25 *cells* in a regular chessboard-like pattern, such that the top row is adjacent to the bottom row, and the left column to the right, as in figure 1. (Hence, this is a torus- or doughnut-shaped world, rather than a sphere.) Every (square) cell on this world has exactly four neighbors (diagonally touching cells are not considered neighbors). In every cell – that is, in every cultural community – one language is spoken. In the initial situation, this is a

1 Using computer simulations or models to investigate aspects of historical linguistics such as language evolution is not a new idea, of course. In fact, over the last decade a growing literature documenting such models has developed. Both the model used here, and the questions it is supposed to answer, differ from what can be found in that literature, however.

different language in every cell. In every cell, the language spoken has a *strength value* between 1 and 1000.

figure 1: 25 cells, each with 4 neighbors



Time is measured in *steps*. In each step all cells get new random strength values (random numbers between 1 and 1000 from www.random.org), and the cell with the lowest value experiences a *language shift*. (In exceptional cases this can happen in multiple cells simultaneously if multiple cells share lowest strength values.) Two things can happen in that cell: either it adopts a language from a neighboring cell, or it develops a new language. (Of course, the model has things backwards – in reality languages disappear *because* they change into something else or are replaced by more powerful ("stronger") languages; they do not disappear and are then replaced. In simulating the process, that difference does not matter, however.) If the strength value of the strongest adjacent cell is higher than a preset *threshold,* then the language of that cell is adopted (which may be the same as the original language of the adopting cell). If it is below the threshold, then the cell develops a new language.

Different results were generated for different sets of random numbers. These different sets of random numbers and their results are identified as "`model #`" below (to be distinguished from "the model", which refers to all of them and their generation procedure together), where `#` is a natural number. Of these `models`, one was randomly selected for presentation and named "`model 1`".

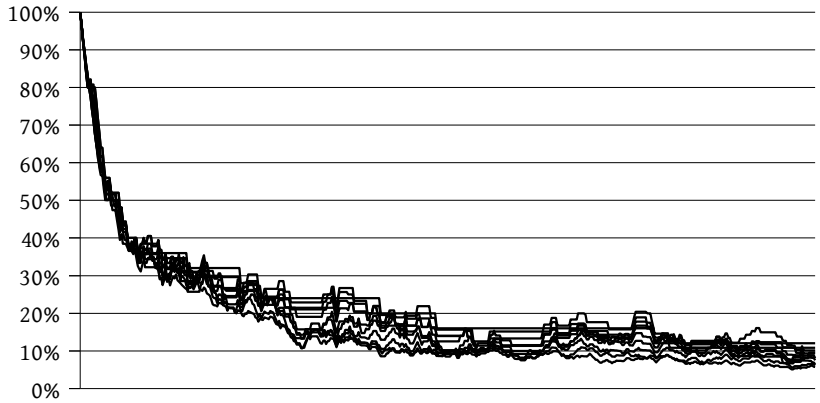
Different thresholds have different effects on the outcomes. Setting it low means that almost always an adjacent language will be adopted and few new languages are developed; setting it high has the opposite effect. Table 1 shows some results for different thresholds for `model 1`. If the threshold is set at 0, then no new languages can be developed. The total number of languages during the period (the column marked "total" on the right) is, therefore, the same as the number of languages in the initial situation: 25. Setting the threshold very high – at 900, for example (bottom row) – results in the development of very many new languages, and very little expansion of languages. In the first case, the number of simultaneously spoken languages gradually decreases (middle columns: 20 after 10 steps, 8 after 100 steps, 3 after 500 steps), while in the latter case, after a small initial decline, it remains stable at around 19 or 20. (It should be noted that in `models 2` and `3` (*i.e.* with other random numbers), at threshold 0, only one language was left after 426 and 213 steps, respectively.)

table 1: threshold effects (`model 1`)

	languages after n steps							DR500
threshold	10	20	50	100	200	500	total	%
0	20	14	9	8	6	3	25	12.0%
400	20	14	10	8	6	4	37	10.8%
500	21	14	10	8	9	5	61	8.2%
550	21	15	10	8	10	7	75	9.3%
600	21	15	11	9	11	9	89	10.1%
650	21	16	14	12	12	11	115	9.6%
700	21	16	14	13	13	11	136	8.1%
750	21	16	14	14	13	13	173	7.5%
800	22	17	15	16	13	14	217	6.5%
850	22	19	17	18	14	17	278	6.1%
900	22	21	19	20	19	20	348	5.7%

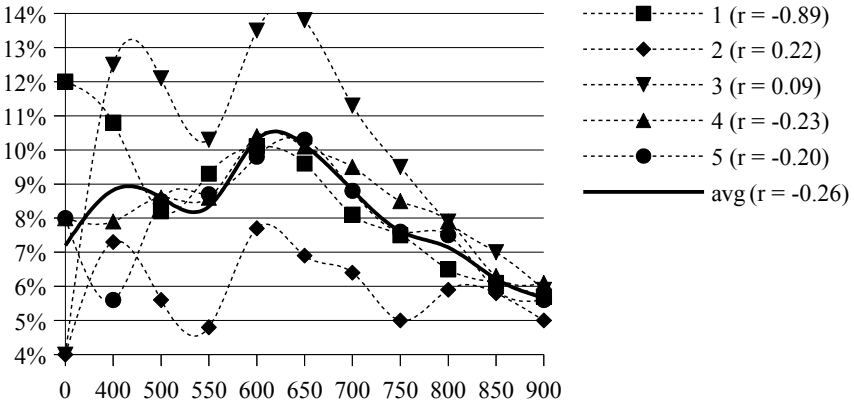
The column on the right marked "DR500" (short for "500 step **D**iversity **R**ate") shows the number of languages spoken after step 500 divided by the total number of languages (the column "total"). Figure 2 shows this percentage (for `model 1`) after any number of steps up to 500 for all 11 different thresholds. Figure 3 shows DR500 (*i.e.* this percentage after 500 steps) for `models 1` to `5` and their average.

figure 2: diversity rates (`model 1`)



Perhaps most surprising about figure 2 is how close together the lines are. Because of technical limitations the lines in the figure are not marked or colored differently, but this does not really matter: different thresholds result in widely different total numbers of languages, but the diversity rates are not far apart, and there is only a weak relation between threshold and level of diversity (and the minimum and maximum values are on different lines at different steps). Figure 3 graphically illustrates this weak relation between threshold and diversity rate. The correlation between the threshold and DR500 in case of `model 1` is -0.89, but in all other `models` it is (much) closer to 0: for most `models`, it is between -0.25 and 0.25. (Correlations for `models 1` to `5` are given in the key to figure 3.)

figure 3: DR500 for 5 `models` and their average (avg), and 11 thresholds



If we ignore the implausible extremes, the model shows that the more probable thresholds result in diversity rates between 7% and 10% after step 500 (see table 1 and figure 3), and that these rates are slowly further converging and decreasing if the model is extended (*i.e.* if more steps are added).

limitations of the model

Before discussing these results and using them and the model itself to answer the questions asked above, it is necessary to assess the limitations of the model and their implications. Most important limitations are related to the very small and fixed number of cells, to the omission of language-internal change, and to the *complete* randomness of strength values.

Increasing the number of cells has two important effects. Firstly, since language shift – substitution of one language for another in the model (noting that the former and later language may be the same in the model) – is not related to the total number of simultaneously spoken languages (unless the total area / number of speakers is too small to sustain so many languages), and one step in the model is defined as a language shift in the weakest cell(s), the time between steps decreases with an increase of the number of cells. This

seems a linear effect: doubling the number of cells, halves the step time. Secondly, increasing the number of cells increases non-direct distances between neighbors. In the model, two adjacent cells have only three cells in between them in the other direction. Hence, even on their non-adjacent sides, they are relatively close to each other. That means that if there is much language expansion (low threshold), increasing the number of cells decreases the change of a cell becoming surrounded and then assimilated by an adjacent language. By implication, with relatively low thresholds, a higher number of cells results in more linguistic variety (in addition of the direct increase as a result of the increase of the numbers of cells). This may be a relatively small effect, however.

If the number of cells does not remain fixed, but either grows under the influence of population increase, or declines under the influence of cultural imperialism, then that does obviously affect the outcomes. In the first case, with growing numbers of cells (cultural communities), the diversity rate drops less and/or less fast, and that effect may be quite significant. Cultural imperialism, on the other hand, may drastically decrease the number of cells (cultural communities) by merging them, thus further reducing the diversity rate.

Change *within* languages may effectively change them into new languages. There are few living languages on earth that stayed the same over longer periods of time. In fact, most languages change so much in a millennium or less, that they really become different languages. This, of course, increases the total number of languages spoken over time, and thus reduces the diversity rate, but – except in case of very low thresholds – the effect may be minor.

In the model, strength values are completely random, but it seems plausible that the strength of a language in a cell at any given point in time is at least partly dependent on its strength before that. This is certainly true if the step time is short, which would be the case with large numbers of cells and languages. Less random strength values would make strong languages stronger

for longer periods of time, giving them more chance to "overpower" neighboring cells. The main effect, therefore, would be a further decrease of the diversity rate. Additionally, strength may also be related to the size of a language, that is, to the number of cells in which a language is spoken, and perhaps to linguistic factors (simplicity, expressive strength, etc.). The model could be adapted to accommodate such effects, and other effects on strength, but this too would merely give an additional advantage to the strongest languages, thus further decreasing the diversity rate.

modeling grammatical diversity

Instead of modeling linguistic change on the level of languages, the model could also be used to simulate variation over time in specific aspects of grammar – different ways of dealing with tense, case, number, and so forth – or to combinations thereof. In such cases, the total number of options would be limited. Regardless of how many ways of marking number or tense there are among existing languages, and how many there could be according to linguistic theories (Universal Grammar particularly), the number is limited by what is *logically* possible: in case of number, noun plurals are unmarked, optionally marked, or obligatory marked, and the latter two may apply either to some or to all nouns, but those five options *seem* to exhaust the possibilities (although it could be argued that differently defined subsets of nouns with or without (optional) plural marking are different options, which would increase the number of options considerably). In addition to the number of different options (*i.e.* languages in the original model), the number of cells in the model may also be smaller: grammatical variation between languages *within* a language family is fairly limited, and what differences there are on that scale are mostly minor variations. Consequently, it may be sufficient to model grammatical diversity on the level of much smaller numbers of language families or (grammatically defined) sub-families.

As a simplified example, consider plural marking and numeral classifiers. These two grammatical features are often considered to be related (*e.g.* Chierchia 1998). Simplifying the above, let us assume that there are three different ways of plural marking: (1) none, (2) optional for most nouns, and (3) obligatory for most nouns; and that the occurrence of numeral classifiers is a dichotomy. This gives 2×3=6 different options. If all of these exist in the starting situation in the model presented above, and the threshold is set at 0, meaning that a variation that disappears cannot return, then before 50 steps the number of variants is reduced to 4, it is down to 3 at around 200 steps, down to 2 at around 300 steps, and only one variety is left after approximately 700 steps. If the threshold is set higher, this changes in an important way: varieties can re-appear. Depending on the threshold, the number of varieties at any step is larger; it is even possible that most or all varieties exist at any given step if the threshold is set high enough. However, unless the threshold is set very high (unrealistically high, because that would imply that grammatical features do not spread, but are always independently re-invented), there is a small and decreasing number of *dominant* varieties, and the other varieties only appear as relatively isolated and minor phenomena.

table 2: variation in plural marking and numeral classifiers

		plural marking		
		none	optional	obligatory
numeral classifiers	no	*pre-Han Chinese*, Udege*	Mongolian, Udege*	**English, French**
	yes	**Japanese, Chinese**	Korean	

For comparison, table 2 shows *some* of the actual variation in plural marking and numeral classifiers. Most languages belong to one of the two cells with boldface text – hence, these are clearly the two dominant varieties – but there

are or were some languages in other cells. A few languages have optional plural marking with or without numeral classifiers. Most interesting of these is Udege, which is placed in two cells and marked with * because in Udege there are optional plural markers but plural marking is ungrammatical if plurality is otherwise marked, for example with a numeral, or "many" (Nikolaeva & Tolskaya 2001). This effectively means that in cases that there *could* be numeral classifiers (which require numerals), there is no plural marking, which would put Udege in the top left cell. That top left cell also contains pre-Han Chinese, which is printed in italics to mark that it is a dead language. The only cell that seems to be empty is the one in the bottom right.[2]

Interestingly, the actual variation, a small number of dominant varieties combined with relatively rare and isolated deviations, looks exactly like what the model predicts at realistic (medium or even medium-high) thresholds. (It would be interesting to test whether a more advanced model including strength numbers representing simplicity and expressive power for the 6 options would also be able to predict *which* options would become dominant.)

estimating the total number of languages throughout history

Taking the above described limitations into account, the model can be used to make a rough estimate of the total number of languages that have been spoken on earth throughout history. This requires an estimation (through simulation) of the relevant diversity rate, which is (primarily) dependent on the number of steps.

Ethnologue (Lewis, Simons & Fennig 2013) records 7,105 languages with at least some speakers, but other estimates are a bit lower (3000 seems to be the lowest, but is probably too low), or occasionally slightly higher (around 8000).

2 This may, of course, be the result of my ignorance: there may be a language that fits in that cell, but that I just never heard of. Although this is not unlikely – I only know a few details about a relatively small number of languages – it may also be possible that this combination indeed does not occur. If strength is partially dependent on simplicity, then this option would be the weakest, giving it a very low chance of survival.

Between half and 90% of living languages may be extinct before the end of the century. *Multitree* lists 1,232 known extinct languages,[3] some of them millennia old, but most relatively recent, and there are, of course, *very* many undocumented extinct languages, even from just a few centuries ago. Based on these figures, it seems plausible that over the last millennium or so, on average *at least* one language disappeared per year, but probably *many* more. To that number, other language changes (such as internal change) must be added to estimate step time. This would result in a very short step time, a few months to half a year at most. However, the world's population has grown exponentially, which may mean that the number of cells (cultural communities) in the model should increase over time as well, and that in turn affects step time in earlier periods: with smaller numbers of cells, step time becomes longer. In other words, "step speed" starts slow and becomes faster and faster. Partially because of that, it is nearly impossible to estimate the *exact* total number of steps; the number may be anywhere between 5000 and 5 times that, and possibly even more. It seems plausible, however, that the number of steps is (much) larger than the number of currently spoken languages – any language spoken now must have changed or be the result of a change *at least* once. That would mean that there have been 7,105 steps at the very least, but probably very many more.

Extending the average of the lines in figure 2 to that many steps results in a percentage somewhere between 1.5 and 3 (or at lower levels of probability, between 1 and 5), and these percentages do not change very much if a *few* thousand steps more are added (but tripling the number of steps reduces the percentage by approximately one third to between 1 and 2). Much more accuracy is impossible because there are too many uncertainties. If we assume that the above mentioned number of languages in Ethnologue is approximately correct, then that would put the total number of languages spoken on earth throughout history between 250,000 and 500,000. If we would be *very*

3 http://multitree.org/codes/extinct.html (accessed October 22, 2013)

conservative, put the diversity rate at 5 instead, and the current number of languages at 3000, the total would be 60,000, but these numbers seem unrealistically "conservative". If we would be less conservative, the number would be *well* over half a million. Considering that most of the limitations mentioned above imply that the actual diversity rate is probably lower than the model suggests, less conservative estimates are most likely more accurate than more conservative ones.

These numbers are quite similar to those mentioned in the introduction. Mark Pagel's (2000) calculations – like those above – depended on some assumptions about the history of language. The total numbers of languages for his 9 models vary between approximately 40,000 and 500,000 (which is very similar to the lower and upper boundaries of 60,000 and 500,000 resulting from the calculations above), but some of Pagel's models are based on relatively implausible assumptions (as is the case with the "very conservative" model above). Ignoring those, the total number would be between 100,000 and 250,000. David Crystal (2002), on the other hand, estimates the total number at between 64,000 and 140,000 (which, in comparison with Pagel's and mine, seems a bit low).

Based on the model, but also taking the other two approaches into account, it seems likely that the total number of languages that has existed throughout history is somewhere around 150,000 (plus or minus 50,000), but *possibly* much higher. This means that the number of languages spoken today is roughly 5% (or less) of all languages that have existed.

possible languages and counterfactual linguistics

Of course, these numbers are rough estimates. Nevertheless, it seems *very* unlikely that there have been less than 100,000 languages throughout human history; more likely the number is much higher. Only a small percentage of those languages is spoken today, and very few are studied extensively. That, of course, does not mean that all variation is lost. It may be the case that

languages that disappeared deviated from known languages only in relatively uninteresting aspects: more small variations of known grammatical patterns, or more minor semantic encoding differences. Given the numbers, it is rather improbable, however, that this is the case. There must be significantly different grammars that have disappeared forever, for example.

What we are left with, is a very small sample of actual historical linguistic variation on earth. That would be no problem if it were an unbiased, *representative* sample, but that – unfortunately – is unlikely too. In the model, strength values are completely random, but in reality they are dependent on political and economical factors, but probably also on linguistic ones – more practical and/or more easily learned languages may be stronger than their opposites. Whatever determines these strength numbers, "strength" determines (or even equals) "reproductive success" and the historical process is to a large extent an evolutionary process: the stronger languages survive, the weaker ones disappear, and by mutation new varieties appear. The result of an evolutionary process, however, is by no means representative of all variety during that process. Current reptile species are by no means representative of all reptile species that ever lived (which includes dinosaurs). Similarly, the collection of currently spoken languages is by no means representative of *all* languages that ever existed. It *may* be, by coincidence, but that is *very* unlikely, and there is no way to know.

Consequently, no statistically significant generalizations about language can be made. On the basis of the small and biased sample of known languages – let alone on the basis of an even smaller subset thereof (*i.e.* the small number of extensively studied and well-documented languages) – nothing can be said with certainty about all languages that existed throughout history, about languages that can or cannot exist, about language in general, or about the philosophical implications of some *apparently* universal linguistic feature. Even if there would be features common to all known languages (aside from features that are necessary to be a (learnable) language, such as some kind of

compositionality), then it may be the case that some lost language(s) did not have that feature. Furthermore, the application of the model to grammar (see above) shows that the actual variation is very similar to what the model predicts, and consequently, apparent universalities *may* just be historical coincidences or the result of a evolutionary (or evolution-like) process. That does not necessarily mean that there *are* no linguistic universals, just that known linguistic variation or lack thereof cannot directly serve as evidence for linguistic universals.

However, it may provide indirect evidence. If, by means of a counterfactual approach to linguistic typology, all logical possibilities with regard to a certain kind of grammatical feature can be determined, then that data – as already suggested above – can be used in a refined version of the model to test Universal Grammar, for example. The model would need more cells and strength values that are partly dependent on prior strengths, dominance (number of cells with that variety), and linguistic aspects such as expressive power and simplicity of a variety. With such (and perhaps some other) adaptations the model could probably develop into a reasonably accurate simulation of grammatical diversity over time. Consequently, if after running that model, a pattern of variation results that significantly deviates from known real variation, then there must be non-historical explanations for that deviation. Depending on the nature of the deviation these would be different explanations, but if the model would predict *significantly* more variation than there is in the real world, then that would be a strong indication for the presence of linguistic universals (with regard to the grammatical feature(s) in the model). If not, then it is more likely that apparent universalities are mere historical accidence, or the result of evolutionary selection.

Many more languages have disappeared than are spoken today, and much linguistic variation is lost. Perhaps, in addition to proving or refuting Universal Grammar, a counterfactual approach to linguistics can recover some of that lost variation. “Language construction” (or “conlanging”) exists as a hobby or

art form (there is no sharp boundary between those two), and there are quite a few languages (called "conlangs") such as Ithkuil and Kēlen[4] that are developed by hobbyists (often with considerable knowledge of linguistics) with just such explorations of what is linguistically possible or impossible in mind. It is not unlikely that many grammatical and other "oddities" that can be conceived, and that could have been part of a learnable and useable language, actually have existed in some language. Testing whether some such "oddity" could be part of a learnable and useable language, which are the minimum requirements for being possible, may require trying it out: placing that oddity in the wider context of a language and using that language, by translating a substantial amount of appropriate text into that language (and back), for example. Thus blurring the line between art and science, constructing languages that significantly deviate from known languages may result in new ideas about what is (im-)possible and what is (im-)probable in linguistic variation, and why.[5]

conclusion

Recapitulating the main findings presented above: the number of dead languages is at least 20 times that of living languages, and the number of languages throughout human history is probably around 150.000 (plus or minus 50,000), but possibly much higher. Furthermore, the approximately 5% thereof that is the set of languages spoken today is the result of a more or less evolutionary process, and is almost certainly *not* a representative sample. This

4 Ithkuil: www.ithkuil.net; Kēlen: http://www.terjemar.net/kelen.php (both accessed on October 22, 2013). Both Ithkuil and Kēlen have – at the time of writing – Wikipedia pages as well.

5 This is not something entirely new. Languages have been constructed by linguists to test linguistic ideas before. Smith and Tsimpli (1995), for example, developed the fictional language *Epun*, which intentionally violates various principles of Universal Grammar and tried to teach it to a linguistic savant. They had relatively little success, which – if there is no other acceptable explanation – may be the strongest evidence for universal grammar to date. However, Bates (1997) rejects most of Smith and Tsimpli's conclusions.

means that actual linguistic diversity is not a reliable source of information about total or possible linguistic diversity.

The question this paper intended to answer was: How much can we learn from *actual* linguistic diversity? The answer: A lot, but not enough to make *universal* claims about language and/or its philosophical implications. Nevertheless, even if actual linguistic diversity cannot be evidence for universal claims and cannot show what is *maximally* possible (with regards to linguistic variation and its implications), it can show us what is *minimally* possible. If different languages involve different ways of representing aspects of reality – such as time, as in the example in the introduction – then studying (and perhaps, constructing) different languages may provide us with different perspectives on reality. Therein lies the real value of studying and comparing languages: in showing us what *can* be, rather than what *must* be. (And the counterfactual approach suggested above may further extend our understanding of language by adding the perspective of what *could* be.)

It should be noted, however, that this is a provisional answer. The model used is a very simple one, with several (noted) limitations. Above, some suggestions for improvement of the model were made.[6] Nevertheless, while these improvements may refine the findings and conclusions of this paper – especially the numbers mentioned – it is very unlikely that this will lead to a different answer to the main question: living languages are a small and biased (unrepresentative) subset of possible and historical (dead plus living) languages. We may be able to learn a lot from the variety among living languages, but not as much as we (or some of us) might like.

6 In addition to those, further ideas for improvement may be found in the extensive literatures on historical linguistics and typology.

references

Bates, E. (1997), "On Language Savants and the Structure of the Mind: a Review of Neil Smith and Ianthi-Maria Tsimpli *The Mind of a Savant*", *International Journal of Bilingualism* 1, pp. 163-179.

Chierchia, G. (1998), "Reference to kinds across languages", *Natural Language Semantics* 6, pp. 339-405.

Crystal, D. (2002), *Language death*, Cambridge University Press.

Davidson, D. (1977), "The Method of Truth in Metaphysics", in: (1984), *Inquiries into Truth and Predication*, Oxford: Clarendon Press.

Evans, N. & S.C. Levinson (2009), "The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science", *Behavioral and Brain Sciences* 32, pp. 429-492.

Lewis, M.P., G.F. Simons & C.D. Fennig (eds.) (2013), *Ethnologue: Languages of the World*, Seventeenth edition, Dallas: SIL International. Online version: www.ethnologue.com.

Nikolaeva, I. & M. Tolskaya (2001), *A Grammar of Udihe*, New York: Mouton de Gruyter.

Pagel, M. (2000), "The history, rate and pattern of world linguistic evolution", in: C. Knight, M. Studdert-Kennedy & J. Hurford, *The Evolutionary Emergence of Language*, Cambridge University Press, pp. 391-416.

Smith, N. & I.M. Tsimpli (1995), *The mind of a savant: language learning and modularity*, Oxford: Blackwell.