Joe McVeigh
Article is on *…And Read All Over* blog at: https://andreadallover.com/2018/3/21/mistakes-will-be-made

Mistakes will be made…

When economists do linguistics.

There's an article coming out in the *Journal of Comparative Economics* called "Talking in the Present, Caring for the Future: Language and Environment" (Mavisakalyan et al. 2018). The authors claim:

- We identify future tense marking in language as a determinant of environmental action.
- Individuals speaking such languages are less likely to behave pro-environmentally.
- Climate change policies are less stringent in places where language marks the future.

This has my Whorfian alarm bells going off like crazy. The language I speak determines how I feel and act towards the environment?! Say what? That is just too bonkers to be true.

Because it's not true. But let's check out why.

If my calculations are correct… you're gonna see some serious shit

I won't get into the language and culture nitty gritty here. Let's just look first at how this article uses languages as a variable, specifically whether a number of languages have future tense marking or not. The most basic thing in an analysis involving language is to accurately describe the language(s) in question. I know how obvious that sounds. But in studies about language, the linguistic analysis needs to be sound. If you don't have that, you ain't got dinky do.

So how's the linguistics in the article by Mavisakalyan et al.? It's… not so good, Al. The authors (Astghik Mavisakalyan, Yashar Tarverdi and Clas Weber) are only partly to blame for the shoddy linguistic work, though, because they take their linguistic analysis from a 2013 paper by Keith Chen called "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets" (link). Chen's paper claims that languages without future tense marking cause their speakers to "save more, retire with more wealth, smoke less, practice safer sex, and [be] less obese". If that sounds bananas, that's because it is.

What the linguistic analysis in the two articles (although it mainly comes from Chen 2013) boils down to is the grammar of how people talk about the future. The languages in the analysis are divided into whether they have "future tense marking" or not. According to Mavisakalyan et al. (2018):

Speakers of languages that lack future tense marking speak about future events in the present tense, i.e. as if they were present. In contrast, languages with future tense marking require speakers to use a distinct form when talking about the future. [p. 3]

The articles classify English as a "strong-FTR" language, or a language which requires a distinct form when talking about the future. Finnish, on the other hand, is classified as a "weak-FTR" language. The distinct form that English speakers are apparently required to use is the modal verb *will*. Mavisakalyan et al. (2018: 5) give this example:

> *Tomorrow they will*$_{auxiliary}$ *drive to Paris*.

First of all, this sentence doesn't sound like it was spoken or written by an English speaker. There is exactly one example of *Tomorrow PRONOUN will drive* in the 520-million-word Corpus of Contemporary

American English (link). There are no hits for this construction in the BNC2014, an 11.5-million-word corpus of spoken British English. GloWbE, a corpus of web-based English, has two hits for *Tomorrow PRONOUN will drive to*. A much more idiomatic example would be *They're driving to Paris tomorrow*. But whatever, English obviously uses the modal verb *will* to refer to the future. Are there other ways English speakers can refer to the future though? Let's check some grammars!

Roads? Where we're going, we don't need roads.

First up is *Navigating English Grammar* (2014, Wiley) by Anna Lobeck and Kristin Denham. They make in interesting point about modals in English:

> **Modals are unique in that their tense interpretation depends on context rather than form. So it isn't really accurate to say, for example, that *can* is the present tense form of the modal and *could* is the past tense form.** In the following sentence, *could* can express past tense, present tense, or future tense, depending on context:
>
> Elizabeth could not leave for Toronto. (yesterday / today / tomorrow)
>
> And *must* in the sentence below expresses either present tense or future tense, again, depending on context [or depending on if there is an accompanying adverb].
>
> Elizabeth *must leave* for Toronto yesterday. (now / tomorrow)
>
> **In these sentences, then, tense is *semantic*, rather than *grammatical* (morphological or syntactic)**. [p. 93; bolding mine]

I guess those example sentences still have distinct forms, but finding out whether they refer to the past, present or future is dependent on either 1) context; or 2) whether they include a time adverbial (*two weeks ago*, *right now*, *in the future*, etc.).

Next we have *A Student's Introduction to English Grammar* by Huddleston and Pullum (2005, CUP). One of the authors, Geoffrey Pullum, has written critically about Chen's (2013) paper on Language Log. This grammar book makes some interesting point about how reference to the future is made in English:

> **The present tense is often used for situations located in future time.** In main clauses this is restricted to cases where it can be assumed that we have present knowledge of a future event, as in:
>
[32]	i	*The next high tide is at 4 o'clock.*	*The sun rises tomorrow at 6.10.*
	ii	*Exams start next week.*	*We arrive home two days before Easter.*
>
> These constructions are called the **futurate**. The future time is usually specified by a time adjunct, marked by double underlining. The two most common cases involve:
>
> - recurrent events in nature whose time can be calculated scientifically (as in [i]);
> - events that are arranged or scheduled in advance (as in [ii]). [p. 45]
>
> […]

> There are some languages that have a three-term tense system contrasting past, present and future. Contrary to what is traditionally assumed, English is not one of them: it has no future tense. It does have several ways of talking about future time, and the most basic one does involve the auxiliary *will*. Nonetheless, *will* belongs grammatically and semantically with the auxiliaries that mark mood rather than with the various markers of tense. [p. 56]

Huh. So English speakers often uses the present tense to talk about the future and the modal verb *will* isn't really a marker of tense. Interesting.

Next up we have the *Longman Student Grammar of Spoken and Written English* by Biber, Conrad and Leech (2002, Pearson). The info in *Longman's* is similar to the previous two (there is no future tense in English, a time adverbial is often used with the present tense to refer to the future), but it also includes this note about another way that people talk about the future:

> Surprisingly, a relatively high proportion of the postmodifiers in conversation are *to*-clauses. Their meaning often points to the future:
>
> *Father's got a lot of <u>things</u>* ***to tell you****.* (CONV) (p. 294)

That's cool. Referring to the future with a *to*-clause. I never thought of that one before.

Finally, there's *Explaining English Grammar* by George Yule (2009, OUP). After showing how *present tense* + *adverb* is often used to refer to the future, Yule (p. 58) explains that "the simple time-line view [or past/present/future] may not be the most helpful guide in making sense of tense and aspect in English […] tense in English is not based on simple distinctions in time." Then Yule explains something that is at the heart of the matter for us. He shows that future events are treated as possibilities through the use of modal verbs and that the modal verb used conveys how strongly the speaker feels toward a future event happening:

> The verb form that is traditionally called 'the future tense' is actually expressed via a modal verb which indicates the relative possibility of an event. This modal also has two forms which convey the closeness (*I* will *live here*) or the remoteness (*I* would *live there*) of some situation being the case, viewed from the situation of utterance. (p. 59)

Notice here that, as in Lobeck and Denham above, Yule doesn't talk about *will* being present tense and *would* being past tense. He also says that the pairs of other modal verbs (*may*/*might*, *can*/*could*) show the same distinction as *will*/*would* in expressing a speaker's belief about how remote or likely a future event is (p. 59).

This idea of possibility and remoteness is interesting. I assume all languages have a way of expressing this. Of course, they don't have to do it the same way that English does. For example, Finnish speakers may commonly use the present tense + time adverbial to refer to the future (just as English speakers do), but Finnish speakers can also convey their feelings of certainty about future events through various means (just as English speakers do). One way is through adding *-isi* to a verb.

> Minä voisin asua siellä / Minä asuisin siellä. = I could live there.

(Note: I am not a native Finnish speaker, nor even a very competent one IMO. If anyone would like to discuss the ins and outs of talking about the future in Finnish, please leave a comment below)

The point is that there are many ways of talking about the future in English:

- By using modals (*can*, *could*, *must*, etc.)
- By using the present tense + a time adverbial
- By letting context do the work
- By using a *to*-clause

But in the Mavisakalyan et al. (2018) and Chen (2013) studies, all of this complexity is boiled down into an either-or, yes-no, ride-or-die distinction between whether each language marks future tense or not. English is considered a language which *does* mark future tense, but as we have seen above, English has no future tense and English only sometimes uses a distinct form when referring to the future. In addition to this, when a distinct form is used, it refers more to mood (or how certain a speaker feels about a future event) than it does to tense. (To be clear, English isn't the only language that Chen gets wrong. According to McWhorter 2014: 95–101, the analyses of Russian, Czech, Slovak, Polish and Korean are also problematic)

There's that word again. "Heavy." Why are things so heavy in the future?

All of this is enough to chuck the analysis out the window, but there's more. Back in 2012, Keith Chen was invited to write a post on Language Log to explain his research (because linguists were hella skeptical). One of the things about Chen's research is that he claims to have adopted the "strong vs. weak FTR language categorization […] from Östen Dahl (and the EUROTYP project)". That means he's claiming that it was a linguistic research project which divided languages into only two categories of if they mark future tense or not. But Östen Dahl, the researcher behind the data that Chen claims to have adopted, showed up in the comments of Chen's LL post and he had words, y'all. Here's the cliffnotes version:

It's a bit scary to see your own analyses invoked in this kind of discussion. But let me note here that the way Keith Chen describes them they sound much more categorical than they are in my own texts. […] I do not specify a binary classification of European languages (let alone the languages of the world) and I do not use the terms "strong-FTR language" and "weak-FTR language". (In the abstract of his working paper, Chen says "what linguists call strong-FTR languages" — what he should have said is "what I call strong-FTR languages" [see the full comment here]

Yikes. Someone turn on a light. It just got real shady in here.

So it looks like Chen might have fudged the numbers – or the letters, as it were – in presenting his linguistic data. When Chen wrote that LL post, he still hadn't published his paper, but he brushed aside the skepticism from linguists by claiming:

> Many linguists may wonder: if the precise function of the EUROTYP classification is not entirely clear, even for deeply studied languages like English, what are we to make of aggregate correlations between this classification and behaviors?
>
> I share this concern, and think it's worth thinking hard about both what it suggests careful work should look like, and how we might interpret results that survive careful analysis. Though I plan to take more steps in the direction of investigating more fine-grained linguistic distinctions in

> future work, **I think it is important to note that despite these issues we can still learn a lot from correlations between even rough typological distinctions and behaviors**. [bolding mine]

He thinks we can learn a lot from oversimplified and probably incorrect linguistic data! Cool! Does economics work like that too? Can I just guesstimate every country's GDP do some economic analysis? 'Cause that would be sweet.

Whatever. Chen published his paper. And that brings us back to where we started. The publication of Chen's paper in 2013 has led Mavisakalyan, Tarverdi and Weber to assume that they can use his classification of languages in their research. They make no attempt to remedy the fact that Chen's classification is a gross and misleading oversimplification of how speakers refer to the future. In fact, despite each of them being speakers of at least one language, they don't seem to have questioned the linguistic analysis at all. They merely state in a footnote that German has the "*potential* to explicitly mark reference to the future" and then move on. I can't believe none of them said, "Wait a minute, I can think of a few ways to refer to the future which don't fit this either-or distinction." Nothing.

This blog post is over 2,200 words at this point and it's on this topic ALONE. And we're not done yet! (Soon, though, I promise)

Now, Biff, don't con me.

There are two more points I want to raise. First, the Chen (2013) and Mavisakalyan et al. (2018) studies find a correlation between the languages supposedly spoken in a country (although Dahl had some words on that one too) and some behavioral traits. Great, but as we have seen the linguistic analysis part of these studies is highly questionable. So where does that leave us? Can we just make up any old claim we want about languages around the world and then look for correlations in economic markers? I suppose we could, but that's not all these studies do. On top of dividing languages into two groups, the studies in question claim that the language are causing people to act in certain ways! This is Bad Linguistics 101.

> "How come you're not saving for the future, Charlie Brown?"
>
> "Because of the way my language *sometimes* marks reference to the future. Ho hum."
>
> #alt-linguistics

Second, there's an elephant in the room that I haven't mentioned. The examples above about referring to the future in English are all pretty much examples of mainstream English. And that presumably is the variety that the studies use. But what about other, non-mainstream varieties? How do African American English speakers refer to the future? What about Chicano English speakers? Or English speakers in Jamaica? Do English speakers in Singapore refer to the future in the same way as the examples above? And English speakers in India? …

And what do we do with the fact that the majority of English speakers are non-native speakers, or the fact that many (most?) people are multilingual? These are relevant questions. We can't just take examples from grammar books which focus on mainstream English and assume that all English speakers everywhere use these forms.

Time to wrap up. I know I've been hard on these papers, but I don't think all economists are bad at linguistics. Heck, the Johnson blog/column in the *Economist* is one of the best out there. And dictionaries lave lexicographers with degrees in economics. I do, however, think that people outside the field of linguistics who are doing research on language need to get their facts straight. And I think that non-linguistic journals need to do a better job refereeing these kinds of studies because they get written up in the press and they apparently lead to more bad studies. Also, readers without knowledge of linguistics will assume these news reports are true.

Everyone grows up speaking a language and people are really good at using language. Those in academia have even learned standard written English. That does not make them language scholars!

References:

Biber, Douglas, Susan Conrad & Geoffrey Leech. 2002. *Longman Student Grammar of Spoken and Written English*. Harlow: Pearson.

Chen, M. Keith. 2013. "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets". *The American Economic Review* 103(2), 690–731. doi:10.1257/aer.l03.2.690

Huddleston, Rodney & Geoffrey Pullum. 2005. *A Student's Introduction to English Grammar*. Cambridge: CUP.

Lobeck, Anna & Kristin Denham. 2014. *Navigating English Grammar: A Guide to Analyzing Real Language*. Oxford: Wiley.

Mavisakalyan, Astghik, Yashar Tarverdi, Clas Weber. 2018. "Talking in the Present, Caring for the Future: Language and Environment". *Journal of Comparative Economics*. doi:10.1016/j.jce.2018.01.003

McWhorter, John. 2014. *The Language Hoax. Why the World Looks the Same in Any Language*. Oxford: OUP.

Yule, George. 2009. *Explaining English Grammar* (Oxford Handbooks for Language Teachers). Oxford: OUP.