Oxford Handbooks Online

The Language-Opinion Connection

Efren Perez

The Oxford Handbook of Polling and Polling Methods *(Forthcoming)*
Edited by Mike Alvarez and Lonna Atkinson

Online Publication Date: Nov 2015

Subject: Political Science, Comparative Politics, Political Behavior
DOI: 10.1093/oxfordhb/9780190213299.013.18

Abstract and Keywords

This chapter critically reflects on the deceptively simple relationship between language and survey response: the language-opinion connection. It discusses what survey researchers actually know about this link, paying close attention to challenges involving conceptualization, measurement, and research design. Throughout, the discussion emphasizes a core theme: despite great advancements in sampling, measurement, and research design, the study of language and survey response is bereft of strong theory. Thus, while the language-opinion connection seems on the surface easy to assess, public opinion researchers have modest theory to explain *how, when*, and among *whom* language influences survey response. Against this backdrop, the chapter outlines several ways forward, stressing in particular the importance of identifying and testing psychological mechanisms.

Keywords: language, survey response, conceptualization, measurement, psychological mechanisms

Introduction

Can language affect survey response? This is a simple question that should be easy to answer. All that is seemingly necessary to test this proposition are measures of people's opinions and the language they interview in. From there, researchers can statistically estimate the association between individuals' reported views and the language of interview (while of course adjusting these estimates for other influences) (Lee 2001; Lee and Pérez 2014; Lien et al. 2004; Welch et al. 1973). In fact, some scholars might wish to assess this relationship experimentally (Druckman et al. 2011; Shadish et al. 2002). Better to fully isolate language's influence via randomization. Only then can one confidently infer language's causal impact on survey response. Simple, right?

Yes, but also utterly deceptive. Because no matter how elegant the research design, survey researchers have yet to answer basic questions about this "language-opinion connection." First, there is a conceptualization issue: identifying which aspects of language influence what facets of survey response (Swoyer 2014). Given that languages vary by their grammatical structure, scholars should specify how such nuances affect people's articulation of opinions, if at all (Slobin 1996). Researchers should also clarify whether language shapes individual opinions by influencing what a person remembers and/or how they remember it (Marian and Neisser 2000). These are not mutually exclusive points, but they are often omitted from existing research on language and survey response (Lee and Pérez 2014; Welch et al. 1973).

Second, there is a measurement problem. In order to claim that interview language affects survey response, scholars must ensure that speakers of different tongues interpret survey questions equivalently (Boroditsky 2001). Individuals who interview in distinct tongues may express dissimilar opinions because they construe survey questions differently (Stegmueller 2011; Pérez 2009; Davidov 2009). Hence, language-opinion gaps might stem from uneven item performance across tongues, rather than language nuances per se.

Third, it is challenging to select a research design to appraise language's influence on survey response. One popular

option is to observationally assess this relationship with national survey data that offer respondents different interview language options (Dutwin and Lopez 2014; Garcia 2009; Lee and Pérez 2014). This approach is strong on representativeness: one can be confident that a correlation between language and opinion(s) reflects a real phenomenon in a mass public. But the same framework is weak on causality: it is impossible to rule out that other (un-)observed traits among individuals, besides language, drive this pattern (Clarke 2005).

Enter experiments, in which interview language can be randomized among bilinguals. This method isolates language, thus allowing tests of its mechanism(s) (cf. Boroditsky 2001; Marian and Kaushanskaya 2007). Yet such experiments can be criticized on several grounds, including, paradoxically, internal validity (i.e., whether language causally impacts survey response) (McDermott 2011). For example, bilinguals often use one of their tongues in specific contexts. In surveys, this manifests as a preference to interview in a certain language (Lee and Pérez 2014). However, if one is randomly assigned to complete a survey in a tongue that is not preferred for that context, this treatment can become reactive in all the wrong ways, such as angering or worrying people, with downstream effects on their opinions (Brader and Marcus 2013).

But the last challenge is the most fundamental. Conceptualization can be addressed, measurement issues can be overcome, and optimal research designs can be selected. Yet all of this is for naught if there is weak theory to explain and anticipate language effects on the varied components of survey response (e.g., recall, reporting) (Tourangeau et al. 2000). There is no silver bullet here. This must be a sustained effort to answer questions such as *how* language shapes survey response, *when* it does so, and among *whom* these effects are strongest. Such an effort might begin by considering whether language differences produce framing effects, or shifts in how "people develop a particular conceptualization of an issue or reorient their thinking about an issue" (Chong and Druckman 2007, 104). It is plausible that by increasing the relative salience of some considerations, language affects how respondents answer survey questions. For example, does interviewing in a tongue that makes explicit gender distinctions (e.g., "his" versus "hers") lead respondents to report less support for policy proposals to tackle gender inequality?

So the simple question about whether language shapes survey response is not so simple after all. In fact, it is not really one question, but several. This means there is plenty of work for survey researchers to do. In order to clarify just how much work there is, I use the next section to pinpoint what scholars know and do not know about the language-opinion connection. I show that while a thick layer of evidence has accumulated on language's association with mass opinion, explicit theories about why this pattern emerges are thin and sparse.

I then explain several findings from cognitive and social psychology, two fields with rich implications for how survey researchers might strengthen their theoretical grip on the mechanism(s) behind the language-opinion connection (Boroditsky and Gaby 2010; Marian and Kaushanskaya 2007; Ogunnaike et al. 2010). I argue that any successful attempt to illuminate language's influence on survey response should consider heeding what psychologists have learned, since most of their insights have not been applied to this domain. The opportunity for synergy across these disciplines is therefore ample and ripe.

Finally, I round out these sections with a detailed discussion of how researchers might make headway on building theories to explain the impact of language on survey response, while addressing issues of conceptualization and measurement along the way. What I say in that section is not meant to be exhaustive. My more modest goal is to highlight what I see as a pressing need to illuminate the microfoundations of language effects on survey response. Let me begin by providing a better sense of what we are up against.

What Survey Researchers (Do Not) Know about the Language-Opinion Connection

With the rise of mass surveys, researchers began assessing public opinion *across* countries, thus encouraging the development of questionnaires in needed tongues (Almond and Verba 1963; Ervin and Bower 1952; Stern 1948). Today, scholars increasingly use multilingual polls to gauge opinions in nations where immigration has brought about population growth (Davidov and Weick 2011; de la Garza et al. 1992; Lien et al. 2004; Tillie et al. 2012; Wong et al. 2011). Take the United States, where Asians and Latinos have arrived in large numbers since 1965. Ryan (2013) reports that about forty million people in the United States speak Chinese or Spanish at home, even though more than 70% of them report speaking English "well" or "very well."[1]

Giving respondents the opportunity to interview in different languages allows researchers to yield more representative portraits of mass opinion (cf. de la Garza et al. 1992; Dutwin and Lopez 2014; Fraga et al. 2010; Hochman and Davidov

2014; Inglehart and Norris 2003; Lien et al. 2004; Tillie et al. 2012; Wong et al. 2011). In many polyglot nations, some people will speak only one language, though it may not be the one used to administer a poll. Some will speak two or more languages, but may prefer to be interviewed in a tongue also not offered by a survey. But others will speak the tongue provided by a poll, although they represent but one stratum in the population. Yet to sample only this last segment because it is easier and cheaper is to misrepresent the opinions of the larger population (Dutwin and Lopez 2014), especially if those preferring to interview in certain tongues display varied attitudes and beliefs (Lee and Pérez 2014). Thus, as societies become (even) more linguistically diverse, the use of multilingual polls will likely continue to grow.

But even as researchers increasingly poll members of mass publics in different tongues, a dense fog hangs over why and how language affects survey response. This ambiguity is reflected in leading explanations about how people articulate opinions (Lodge and Taber 2013; Tourangeau et al. 2000; Zaller 1992). These frameworks suggest that survey response depends on the question being asked and the considerations it evokes (Zaller 1992). Specifically, survey questions activate concepts in long-term memory, which is associatively organized (Lodge and Taber 2013). This means concepts are linked to each other in a lattice-like network, in which stimulation of one energizes others via spreading activation (Collins and Loftus 1975). Once relevant concepts are aroused, they are recruited from long-term memory into working memory—the "top of the head"—where one assembles them into a response (Zaller 1992). Yet nowhere in these theoretical accounts does language explicitly play a role.

This omission is at odds with what some survey researchers are finding. Several studies show that public opinion is reliably associated with interview language (Lien et al. 2004; Pérez 2011; Welch et al. 1973). Lee (2001) reports robust correlations between interview language and opinions on several topics in the Latino National Political Survey (LNPS; 1988–1989), a seminal study of U.S. Latinos. Lee and Pérez (2014) reveal that such patterns also emerge in newer data sets, like the Latino National Survey (LNS; 2006). For example, LNS respondents interviewing in English report 10% more knowledge about U.S. politics than those interviewing in Spanish. Moreover, Garcia (2009) finds that about a fifth of LNS respondents changed interview languages—from English to Spanish or Spanish to English—with this switching affecting people's opinion reports.

These associations between individual opinions and language of interview are generally robust to statistical controls and reproducible across several data sets and different populations that are linguistically diverse (e.g., Asian Americans) (cf. Lee 2001; Lee and Pérez 2014; Lien et al. 2004). Yet their interpretation remains open to debate—and for precisely some of the reasons I discussed in the introduction. Let us start with the question of research design.

Correlations, Correlations, Correlations

Most evidence affirming a language-opinion connection derives from correlational studies of survey data that are representative of populations like Asian Americans or Latinos (Garcia 2009; Lee 2001; Lee and Pérez 2014; Lien et al. 2004; Welch et al. 1973). Finding that individual opinions are correlated with interview language is remarkable, because it implies that what survey respondents report is shaped by the tongue they use to complete a poll. But the correlational nature of these studies raises strong concerns about omitted variable bias (Clarke 2005), since interview language is self-selected by respondents, not randomly assigned. Scholars have dealt with this by adjusting estimates of language effects for a litany of observed covariates (e.g., age, education, language proficiency) (Garcia 2009; Lee and Pérez 2014; Welch et al. 1973). But this ignores unobserved differences between respondents and makes the generated results increasingly model dependent (Clarke 2005). Clearer and stronger evidence, then, is needed to bolster the claim that language independently influences survey response.

Apples and Oranges

A thornier issue involves measurement: specifically, ensuring that speakers of different tongues sense the same reality. This is formally known as measurement equivalence, or what Horn and McArdle (1992, 117) refer to as "whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute" (Davidov 2009; Harkness et al. 2003; Jacobson et al. 1960; Pérez 2009). Applied to the case of language and survey response, measurement equivalence is achieved if survey questions capture the same attitude, belief, value, and so forth from respondents who interview in different languages. Consider the assessment of group identity in a survey. Attaining measurement equivalence here demands that items appraising this construct do, in fact, capture the same form of identity, to the same degree, across respondents interviewed in different tongues. If these conditions are not met,

scholars risk comparing "apples" to "oranges" (Stegmueller 2011).[2]

Despite painstaking questionnaire translations, however, speakers of varied tongues often interpret survey items differently (Harkness et al. 2003; Pérez 2009; Stern 1948). Survey questions aim to measure opinions that are latent and not directly observed. This means a person's *observed* score (y_i) on a survey question is conditional on their true opinion score (η) and nothing else. When $F(y_i \mid \eta)$ holds, observed differences in answering a question reflect true opinion differences. But if speakers of varied tongues interpret a survey item differently, a person's response to a question is conditional on his or her opinion and language group (g_i)—that is, $F(y_i \mid \eta, g_i)$. [3]

When this happens, language-opinion differences are conflated with item quality differences, making it harder to pin a causal effect to language (Stegmueller 2011). Moreover, if questions are more difficult for some language speakers, then they will misreport their opinion level. Pérez (2011) shows that even at equal levels of political knowledge, Spanish interviewees were less likely than English interviewees to correctly report which candidate won the most votes in their state in the 2004 presidential election, due to item bias. Similar results arise in items measuring other traits, with item bias yielding "false positives" in sample survey data. More reassurance is thus needed that any language-opinion gap is real rather than a measurement artifact.[4]

Where's the (Theoretical) Beef?

But even if the preceding methodological challenges are resolved, there is the issue of theory—or rather, a relative lack of it. Most research on the language-opinion connection focuses on detecting this relationship and ruling out alternative influences. Less emphasis is placed on why language is even linked to survey response in the first place (Garcia 2009; Lien et al. 2004; Welch et al. 1973; Wong et al. 2011). For example, Lee and Pérez argue that language gaps "cannot be *reduced to* a technical matter about omitted variable bias, measurement error, or status deference" (Lee and Pérez 2014, 20). But studies like these neglect to clarify *how* language shapes *which aspect of* survey response. Hence, a more convincing case still needs to be made about the pathway(s) linking language nuances to individual differences in survey response.

So, evidence on the language-opinion connection is assailable on several fronts. Yet my sense is that these challenges can be conquered by looking beyond established results in public opinion research. One area worthy of attention is the work of psychologists, which illuminates the micromechanisms behind language effects on thinking. Indeed, if our target is to develop more agile theories to explain the language-opinion connection, then heeding these psychological insights stands to make survey researchers sharper marksmen. Let me explain why.

Language and Thinking: The View from Psychology

The idea that language affects people's thinking is often associated with the amateur linguist Benjamin Whorf, who (in)famously claimed that people's thoughts are completely determined by language (i.e., linguistic determinism) (Swoyer 2014; Boroditsky et al. 2003). In one of his stronger expositions of this view, Whorf (1956, 221) asserted that users of varied tongues are led to "different … observations … and … evaluations of externally similar acts …, and hence are not equivalent as observers but must arrive at … different views of the world."

While certainly intriguing and ambitious, Whorf's hypothesis slowly ran aground on several shoals of criticism, eventually sinking his outlook on language and thinking. Some of the distress experienced by his hypothesis was self-inflicted. For all of its bravado, Whorf's contention was remarkably light on evidence, with most support based on anecdote and personal observation of the languages he studied. Some of the trouble, though, arose from unpersuaded skeptics, who marshalled evidence that shredded the hull of Whorf's hypothesis. Especially damaging here were studies showing that non-English speakers could learn English color categories, even though their native tongue had few words for color (Heider 1972; Rosch 1975).

Consequently, many scholars have found Whorf's hypothesis untenable and unfalsifiable (Boroditsky 2003; Swoyer 2014). But a new generation of psychologists has refashioned his claim into weaker, but still interesting and testable, versions (Boroditsky 2001; Fuhrman et al. 2011; Marian and Neisser 2000). These researchers have threaded together varied theoretical accounts about language's cognitive effects, with their findings yielding a rich tapestry of evidence. Let us examine some of the parts making up this whole.

Grammatical Differences and "Thinking for Speaking"

One way psychologists have rehabilitated Whorf's hypothesis is by drawing on Slobin's notion of "thinking for speaking." Slobin (1996) argues that languages vary in their grammatical organization, which obliges speakers to focus on varied aspects of their experience when using a given tongue. As he explains, the "world does not present 'events' and 'situations' to be encoded in language. Rather, experiences are filtered through language into verbalized events." For example, *gender-less* languages, like Finnish, do not require speakers to designate the gender of objects. In fact, even the word for "he" and "she" is the same in these tongues. In sharp contrast, *gendered* tongues, like Spanish, require speakers to differentiate genders and assign it to objects. For example, to say that "the sun is rising," Spanish speakers must denote the masculinity of the "sun" by using the definite article *el*, as in "*el sol va saliendo*."

Using this framework, psychologists have gathered new and more convincing evidence that language can affect various aspects of cognition, including how people represent objects in memory (Boroditsky et al. 2003) and how they distinguish between shapes and substances (Lucy and Gaskins 2001). One research stream studies how quirks of grammar yield nuances in "thinking for speaking" and thus, variations in how people sense or judge phenomena (Boroditsky and Gaby 2010; Boroditsky et al. 2003; Cubelli et al. 2011; Fuhrman et al. 2011; Vigliocco et al. 2005). Here Boroditsky (2001) teaches us that languages, like English and Mandarin, vary by how they conceptualize time. English speakers delineate time horizontally with front/back terms, as in "what lies *ahead* of us" and "that which is *behind* us." Mandarin speakers employ front-back terms, too, but they also use vertical metaphors, as in earlier events being "up" and later events being "down."

Such differences should hardly matter, right? Yet careful research shows that these language differences can become important when individuals think about time. For example, Boroditsky (2001) primed English and Mandarin speakers with horizontal cues (e.g., a visual of a black worm ahead of a white worm) or vertical ones (e.g., a visual of a black ball above a white ball). Remarkably, she found that Mandarin speakers were milliseconds faster in confirming that March precedes April when primed vertically rather than horizontally.

Other scholars have shown that "thinking for speaking" affects people's sense of spatial locations (Li and Gleitman 2002). For example, Dutch and Tzeltal are tongues that describe spatial relations in relative and absolute terms, respectively. Seizing this nuance, Levinson (1996) sat Dutch and Tzeltal speakers at a table with an arrow pointing right (north) or left (south). He then rotated subjects 180 degrees to a new table with arrows pointing left (north) and right (south), asking them to choose the arrow that was like the earlier one. Dutch speakers generally chose in relative terms. If the first arrow pointed right (north), then they chose the arrow that pointed right (south). In contrast, Tzeltal speakers generally chose in absolute terms. If the first arrow pointed north (right), then they chose an arrow that pointed north (left).

Language and the Encoding Specificity Principle

The studies discussed above powerfully illustrate how "thinking for speaking" can clarify the influence of language on cognition. But this is not the only way to explain language's influence over people's minds. Other research has drawn inspiration from what psychologists call the *encoding specificity principle*, the idea that people recall information more easily when there is a match between how they learned it (i.e., encoding) and how they draw it from memory (i.e., retrieval) (Tulving and Thomson 1973; cf. Godden and Baddeley 1975; Grant et al. 1998).

Accordingly, Marian and associates argue that language facilitates memory recall when the tongue used to retrieve information (e.g., childhood memories) matches the tongue in which the content was acquired (Marian and Neisser 2000). For example, Marian and Fausey (2006) taught Spanish-English bilinguals information about history, biology, chemistry, and mythology in both tongues. Subjects' memories were more accurate, and their recall faster, when they retrieved the material in the language they learned it in. Similarly, Marian and Kaushanskaya (2007) asked Mandarin-English bilinguals to "name a statue of someone standing with a raised arm while looking into the distance." Subjects were more likely to say the Statue of Liberty when cued in English, but more likely to identify the Statue of Mao Zedong if cued in Mandarin.

Rounding out this research, Marian and her colleagues have also demonstrated that memories encoded in a specific language are more emotionally intense when retrieved in that tongue. Marian and Kaushanskaya (2004) asked Russian-English bilinguals to narrate a life event that came to mind when given a prompt, with the researchers tape-recording all narrations. Two raters coded all the narrations for their emotional intensity. In line with the encoding specificity principle,

the authors found that subjects articulated narratives that were more emotionally intense when the language of encoding was congruent with the language of retrieval.

The Interface Between Language and Culture

Another fruitful research area examines the bond between language and culture. Social psychologists have found a strong link between varied tongues and specific cultures, in which any "two languages are often associated with two different cultural systems" (Hong et al. 2000, 717; cf. Bond and Yang 1982; Ralston et al. 1995). The paradigmatic example is research on the *private* and *collective* self (Triandis 1989). This work suggests a *private* and *collective* self exists in all of us, with the former revealed in thoughts about the individual person (e.g., "I am great") and the latter in thoughts about a person's group membership(s) (e.g., "I am a family member") (Triandis 1989). Yet the relative emphasis a person places on these selves varies between cultures, with people in individualist cultures like the United States reporting more *private* self-cognitions than peers in collectivist cultures like China (Trafimow and Smith 1998; Trafimow et al. 1991). For example, Ross and colleagues (2002) randomly assigned Chinese-born subjects in Canada to complete a study in Chinese or English. Revealingly, subjects who participated in Chinese reported more cognitions about the self in relation to others ("I am a family member") than did those participating in English.

The Automatic Influence of Language on Thought

Finally, within this sea of studies on language and thought there is an isle of work suggesting that language automatically shapes people's attitudes (Danziger and Ward 2010; Ogunnaike et al. 2010). What makes this research compelling is that the attitudes people express in these studies are not self-reported, but implicit—that is, nonverbalized, spontaneously activated, and difficult to control (Pérez 2013). This implies that language's cognitive influence is sparked well before people start to cobble together an opinion to report (Lodge and Taber 2013).

Danziger and Ward (2010), for example, had Arab Israeli undergraduate students complete an Implicit Association Test (IAT), a millisecond measure that assesses how quickly people associate different objects like racial groups with words of varied positive or negative valence (Pérez 2013). The IAT here measured automatic associations between Arabs (Jews) and words with negative (positive) valence. Subjects completed the IAT in either Arabic or Hebrew on a random basis. Strikingly, Arab-Israeli bilinguals evaluated Arabs less positively when completing the IAT in Hebrew than in Arabic. Yes, you read that right: people's spontaneous judgment of ethnic groups *shifted* with the language used to evaluate them.

This tantalizing result does not seem to be a fluke, for other researchers have detected a similar pattern, not once, but twice—and in different samples, no less. In a study of Moroccan Arabic-French bilinguals, Ogunnaike and associates (2010) found that subjects automatically evaluated Arabic names more positively than French names when completing an IAT in Arabic. Not to be outdone, a second study revealed that U.S. Spanish-English bilinguals automatically evaluated Spanish names more positively than English names when completing an IAT in Spanish.

These studies are also crucial for another reason. We learned earlier that comparing the opinions of varied language speakers is difficult because people may construe survey questions differently. One solution is to establish language effects on nonlinguistic tasks (Boroditsky 2001), which do not require the use of language (or very little of it). By showing language effects on the IAT, in which no verbal expression of attitudes occurs, Danziger and Ward (2010) and Ogunnaike and colleagues (2010) bolster the claim that language yields nuances in people's thinking.

Toward the Psychology of Language Effects on Survey Response

Clearly, cognitive and social psychologists have amassed a trove of theoretical insights, complete with empirical evidence, about how language can affect people's thinking. But is any of this relevant for survey response? I would like to think so, but the situation is a little more complex than that. First, most of the evidence we just discussed is from small-scale experiments ($N < 50$) with convenience samples (Boroditsky 2001; Cubelli et al. 2011; Fuhrman et al. 2011; Marian and Neisser 2000). Low statistical power thus becomes a concern. With so few observations, the deck is stacked against finding a true effect in these tiny samples; and, when an effect is detected, the likelihood that it is real and not due to chance is worryingly low (Button et al. 2014; Cohen 1992).

Second, these challenges are compounded by the "college sophomore" problem (Sears 1986). Most studies of

language effects center on undergraduate college students, which raises concerns about external validity or whether language can influence thinking across different subjects, research settings, timings, treatments, and outcomes (McDermott 2011; Shadish et al. 2002). College students are a thin slice of any population, which is a problem insofar as scholars wish to make claims about whether language affects survey response in the *mass public*, where the public entails more than just "college sophomores." Thus, one way to increase the external validity of language effects research is to conduct experimental tests in nonlab settings, with more variegated samples, and with survey response as a dependent variable—in other words, in a public opinion survey.

Third, there is a tangled knot between language and culture. Those who do studies on language and thinking find it difficult to rule out that the main driver of observed differences between varied language speakers is the tongues they use, not the cultures they inhabit (Bond and Yang 1982; Ralston et al. 1995; Ross et al. 2002; Trafimow et al. 1991). An even bigger specter, perhaps, is that language might be endogenous to culture, which would make it hard to sustain the claim that language causes shifts in people's survey reports (King et al. 1994).

These are all delicate issues that complicate the wholesale transfer of psychological insights to the realm of survey response. But they are not insurmountable, and they should not detract from formulating theories to explain the language-opinion connection. For example, low statistical power is easy to "fix." Increasing any study's power simply demands that researchers be more explicit about the effect sizes they anticipate a priori, while collecting enough observations to be able to detect effects of that size if they do, in fact, exist.

Public opinion researchers can also credibly address the "college sophomore" issue, though the solution is not as obvious as it might seem. If the problem is that most studies of language effects are based on students, then the temptation is to run such experiments on samples that are representative of a population of interest. But the broader issue here is external validity: the extent to which a language-opinion connection arises, not just in larger and more heterogeneous samples, but also across varied research settings, timings, treatments, and outcomes (McDermott 2011; Shadish et al. 2002). For the language-opinion connection, this entails answering basic questions, such as whether language shapes survey response across varied samples and data-collection modes (i.e., online, telephone, and face-to-face surveys).

Ruling out that language is endogenous to culture can also be overcome with heavy conceptual lifting. "Culture" is a loaded term that means different things to different people. Hong and colleagues (2000, 710) note that a common but static view of culture defines it as a "highly general structure, such as an overall mentality, worldview, or value orientation." Yet a more dynamic view of culture deems it a shared mental map that includes "unstated assumptions, tools, norms, values, habits about sampling the environment, and the like" (Triandis and Suh 2002: 136), which can be activated by speaking a specific tongue (Ross et al. 2002; Trafimow et al. 1991). If one views culture statically, then distinguishing it from the tongue one speaks will involve manipulating language across distinct cultures and observing its effect on similar outcomes, which would reveal whether language comparably affects thinking in varied cultural settings (McDermott 2011). But if one views culture fluidly, the influence of language on it does not need disentangling, since language is a trigger to cultural knowledge. It all depends on one's perspective.

Ultimately, however, resolving these challenges only clears the path for the more difficult task that is theory building. In particular, public opinion researchers who are interested in language effects must still clarify how, when, and among whom survey response is affected by the tongue in which individuals interview.

How Does Language Affect Survey Response?

The most fundamental question to answer, as I see it, concerns how the tongue one speaks influences survey response. This is a lot more difficult than it seems, because it requires researchers to specify what aspect of survey response is affected by language. For instance, does language affect the *content* of people's attitudes, beliefs, and values? Does it affect how those considerations are *retrieved*? Or does it influence how they are *expressed*?

One promising avenue to pursue is to draw explicitly on Slobin's (1996) notion of "thinking for speaking." This is the idea that languages vary in their grammatical organization, which obliges speakers to focus on different aspects of their experience when using a given tongue. As Slobin (1996, 75) explains, this is "the thinking that is carried out, on-line, in the process of speaking." It is the act of encountering the contents of the mind in a way that is consistent with the grammatical demands of one's tongue. The trick here, then, is to precisely identify how such quirks of language can

affect survey response.

That grammar might shape survey responses is not farfetched. For example, Pérez and Tavits (2015) study the grammatical nuances between gendered and gender-less languages to study public attitudes toward gender inequality. They argue that speaking a *gender-less* tongue promotes gender equity by failing to distinguish between male and female objects. Speakers of a gender-less language should thus find it harder to perceive a "natural" asymmetry between the sexes, which leads them to be more supportive of efforts to combat gender inequality.

To test this, Pérez and Tavits (2015) randomly assign the interview language in a survey of Estonian-Russian bilingual adults in Estonia, in which Estonian is a *gender-less* language and Russian a *gendered* tongue. Compared to Russian interviewees, Estonian interviewees are more likely to support making family leave policy flexible so that a father can stay home with a baby. They are also more likely to endorse efforts to recruit more women to top government posts and the nomination of a female defense minister. Across these outcomes, the boost in the probability of support induced by interviewing in Estonian ranges between 6% and 8%, which is noteworthy because all other differences between bilinguals are held constant via randomization. Further, these authors rule out that support for efforts to combat gender inequality do not come at men's expense, because gender-less language speakers become either pro-female or anti-male.

Yet not all public policy proposals break down along gender lines, so perhaps "thinking for speaking" has limited applicability beyond this crucial, but narrow, domain. But recall that "thinking for speaking" variations arise in other areas, like conceptions of time and space (Boroditsky 2001; Boroditsky and Gaby 2010), which are incredibly important for how the public evaluates policy proposals. Let me illustrate with temporal conceptualizations.

Some tongues differ by how future oriented they are. Chen (2013) explains that languages vary in the degree to which they dissociate the future from the present. Tongues with a strong future-time reference (FTR) crisply distinguish the future from the present, while weak FTR languages equate the future and present. Chen (2013) argues that weak-FTR languages should lead people to engage more in future-oriented behaviors, because those tongues conflate "today" with "tomorrow," finding that speakers of weak-FTR tongues save more, retire with more wealth, smoke less, practice safer sex, and are less obese.

But how might such insights explain people's policy attitudes? One possibility acknowledges that time horizons play a major role, as evidenced by research on the temporal dynamics of public opinion (Gelman and King 1993; Stimson 2004). Language nuances in time perception could plausibly affect public support for policies with long-run consequences, such as ones addressing climate change (Pérez and Tavits n.d.; Villar and Krosnick 2011). Net of one's ideology or attention to the issue, support for such policies might be weaker among speakers of tongues with a strong FTR, since they can more easily discount the future, when climate change consequences will be more pressing than they are now.

The same is true of public support for changes to entitlement programs (e.g., pensions, health insurance). Mass publics in many nations face the prospect of reforming expensive entitlement programs *today*, so that their governments can remain solvent *tomorrow* (Pew Research Center 2014). But perhaps to people who speak a tongue that allows them to more easily brush off the future, government insolvency does not feel like an immediate problem. Thus, public resistance to such reforms might partly arise from language, with speakers of strong FTR tongues expressing less support, since it is easier for them to downplay the future.

Of course, these last two examples offer more promise than fact. Yet I highlight them to illustrate how "thinking for speaking" can help public opinion researchers assess not only whether language can affect survey response, but also in which domains.

When Does Language Affect Survey Response?

Another useful question to consider is when language impacts survey response. One way to do this is by pushing on the boundaries of what we already know about this phenomenon in a world where language *does not* seem to matter. There, people express opinions on the basis of considerations evoked by survey questions (Zaller 1992). Think of framing effects in which simple changes in the phrasing of survey items generate noticeable changes in people's responses (Chong and Druckman 2007). Smith (1987), for example, shows that survey respondents are much more supportive of spending on "assistance for the poor" than on "welfare." That basic word changes affect individual preferences, by evoking varied considerations, implies that people's opinions might be shaped by the very language they use to report

those opinions. After all, Marian and colleagues suggest that individual recall of information is facilitated when the tongue used to retrieve a memory matches the tongue in which a memory was learned (Marian and Kaushanskaya 2007; Marian and Fausey 2006; Marian and Neisser 2000).

Drawing on Marian and associates' insights, Pérez (2014) argues that political concepts, such as U.S. political facts, are more associated with some languages (e.g., English) than others (e.g., Spanish). Hence, some political concepts will be more mentally accessible on the basis of interview language. Randomizing the language of an interview among a national sample of English-Spanish bilingual Latino adults ($N = 530$), Pérez (2014) finds, inter alia, that English interviewees report up to 8% more political knowledge than Spanish interviewees. That is, just interviewing in English allows people to report more knowledge about American politics, because those facts are more strongly tied to English. By the same token, English interviewees report reliably lower levels of national origin identity (e.g., "I am pleased to be Mexican"), since the nation of origin is a concept that is more strongly tied to Spanish.

Pérez (2014) then buttresses these results in three ways. First, he analyzes his survey items to establish that such language gaps are not measurement artifacts (i.e., multigroup confirmatory factor analysis) (Davidov 2009; Pérez 2009; Stegmueller 2011). Second, he shows that these language-opinion gaps are not mediated by bilinguals experiencing strong emotional reactions (i.e., anxiety, anger, and pride) to interviewing in one of their tongues (Brader and Marcus 2013). Third, he demonstrates that opinion differences by language do not stem from English interviewees feeling more efficacious by interviewing in a dominant tongue, which would motivate them to more thoroughly search their memories for relevant content.

Nevertheless, Pérez's (2014) insights stem from an online survey experiment. True, opinion data are increasingly gathered on the Web, but increasingly is not the same as always. Many researchers still assess opinions via telephone, face-to-face, and mixed designs (Dutwin and Lopez 2014; Fraga et al. 2010; Wong et al. 2011), and what analysts find in online polls is unlikely to wholly transfer to phone or in-person surveys. For example, online polls are anonymous compared to phone or in-person surveys, which can affect the prevalence of reported attitudes and behaviors (e.g., Piston 2010). Once scholars veer into contexts in which interviewees interact with live interviewers on the phone or face-to-face, the relative anonymity of online surveys is replaced with interpersonal pressures arising from respondents communicating their opinions to an actual person. With live interviewers, it is plausible that respondents will use a survey to "prove" their skill as a speaker of the interview language, perhaps especially when the interviewer is a member their own race/ethnicity (Davis 1997). Alternatively, respondents might use a survey context to show they are *more* skilled than the interviewer in the language of survey response; again, perhaps especially when a respondent and interviewer share the same race/ethnicity.[5]

Scholars can also exploit survey mode differences to shed light on when language effects are independent of culture (Swoyer 2014). To clarify this, one can imagine manipulating respondents' interview language *and* their assignment to an online or face-to-face survey. The assumption here is that if a survey context shifts from an anonymous online setting to a face-to-face context, the pressure to adhere to cultural norms strengthens, because one is directly observed by an interviewer. If the language-opinion connection is independent of culture, one should observe reliable opinion differences by interview language, with small differences between survey modes.

Finally, researchers can further explain when language affects survey response by clarifying how the tongue one speaks maps onto specific domains. Recall that Pérez and Tavits (2015) argue that interviewing in a nongendered tongue (i.e., Estonian) liberalizes one's views about gender inequality. However, they also show this effect is less likely when strong social norms surround a topic (e.g., people should disagree that "men are better political leaders than women"). In the absence of strong norms, language has a wider berth to affect survey response. Scholars can extend this insight by ascertaining whether the language-opinion connection also depends on how crystallized one's attitudes are, with less crystallized attitudes being more malleable. Here Zaller (1992) and others (Lodge and Taber 2013; Tourangeau et al. 2000) remind us that individuals do not possess ready-made opinions on many matters, leading people to often report opinions formed on the basis of accessible considerations. Language effects might therefore be more likely when one's opinion on a topic is not preformed.

Whose Survey Response Is Affected by Language?

Most research on language's influence on cognition focuses on average treatment effects, that is, on whether nuances between tongues causally impact an evaluation or judgment (cf. Boroditsky 2001; Marian and Neisser 2000; Lee and

Pérez 2014; Ross et al. 2002). Less explored is whether such language effects are heterogeneous, which demands the identification of moderators and their integration into research designs.

At least two possibilities come to mind. The first is cognitive sophistication, a workhorse of public opinion research (Delli Carpini and Keeter 1996; Luskin 1987). Sophisticated persons possess more and better organized attitudes and beliefs—all considerations that they are more adept at tying to their judgments. Language-opinion gaps might thus widen across sophistication levels, because experts might be more skilled at "thinking for speaking" (Slobin 1996) or smoother at retrieving relevant considerations (Marian and Neisser 2000). Such possibilities can be tested by measuring sophistication levels and entering them as a moderator in observational/experimental analyses, or by blocking experiments on their basis. Either way, a clearer sense of where scholars are most likely to uncover language effects should emerge.

Another possible moderator draws on the immigrant origins of many bilingual communities: generational status. This attribute reflects how far removed one is from the immigrant experience (Abrajano and Alvarez 2010; Portes and Rumbaut 2006). First-generation individuals are foreign born. Second-generation individuals are born in a host society to foreign-born parents. Members of the third generation or later are born in a host society to native-born parents. Seizing on this characteristic, one might reason that the accessibility of American identity increases among later generation individuals, who are more likely to speak English. Since American identity is conceptually associated with the English language (Pérez 2014), interviewing in English should make this identity more accessible across generational status, thereby producing a gap in American identity levels within immigrant groups.

The question of whose opinions are swayed by language differences can also be answered by tinkering with the bilingual populations that are studied. Not all bilinguals are created equal. For example, among U.S. Latinos, bilinguals typically speak English and Spanish. But some of these individuals learn Spanish first, and then English, whereas others will learn both languages in the opposite sequence. Hence, the order in which bilinguals learn their languages, and their standing preference for one of them, might affect the strength of language effects. I stress, however, that there is no "perfect" sample of bilinguals. Instead, heterogeneity in bilinguals' language repertoires might be usefully exploited to establish boundary conditions for language effects. That is, among what types of bilinguals do we (not) find language effects?

These conditions can be probed by considering how degrees of bilingualism among self-reported bilinguals qualify language effects. Bilinguals are often identified through self-reports of skill in two languages (e.g., "Would you say you can read a newspaper or book in Spanish [English]?"). But this approach lends itself to slippage: people may (un)intentionally misreport their level of skill in two languages. Determining the reliability of the language-opinion connection will ultimately depend on whether scholars can consistently uncover it *across* studies whose subjects' degree of bilingualism varies. Yet before we get to that chain of studies, single investigations will be the order of the day. Figuring out how reliable the language-opinion connection is in single studies will require scholars to validate the self-reported data they collect from bilinguals. One way is to gauge attitudes with multiple items so that measurement error can be diagnosed, with lower degrees of "noise" validating the self-reported data.

With so much emphasis on bilinguals, it easy to forget that language effects also imply an influence on monolinguals. Acknowledging this can help scholars make better sense of puzzling empirical patterns in public opinion research. For example, why is it that in the United States, Latinos report substantially lower levels of knowledge about American politics, even after holding constant individual differences in established correlates of political information (e.g., age, education, political interest)? Well, if facts about U.S. politics are generally more associated with the English language (Pérez 2014), then for the many Latinos who prefer to interview in Spanish, this information will be systematically less accessible, thus contributing to the observed deficit in Latino knowledge about U.S. politics.

Methodologically, researchers can gain a better grip on language's influence on survey response by actively incorporating monolinguals into experimental designs. One way is for researchers to employ monolinguals as something of a control group, allowing scholars to make better sense of language effects (Ross et al. 2002). Here researchers can use monolinguals to see how the opinions of bilinguals from the same culture compare. For example, are the opinions of Latino bilinguals who interview in English comparable to those of Latinos who are English monolinguals? Researchers might also consider using monolinguals from different cultures, such as whites who are English monolinguals, and compare them to Latino bilinguals who interview in English. If the opinions of the former resemble those of the latter, then it is harder to say that culture drives opinion differences.

Finally, most psychological and political science studies of language effects investigate differences between individuals,

usually bilinguals within specific national contexts. But bilinguals are a unique subpopulation, which calls into question the generalizability of such results to a larger context beyond their specific setting. One way to further validate these results is by analyzing cross-national differences in the language-opinion connection. This can involve examining the impact of aggregate language differences on aggregate indicators of opinion. It can also entail analyzing differences between individuals from nations that primarily speak different tongues. Finding further evidence like this (Chen 2013) can bolster the case that observed language effects are not a strict function of the within-nation analysis of bilinguals usually undertaken by researchers.

Conclusion: So What, and What to Do?

The preceding pages underscore that failure to include language in models of survey response risks distorting our conceptual understanding about how people form opinions, since language can affect what is activated in people's minds, what people retrieve from memory, and what individuals ultimately report in surveys. But some might be tempted to ask: So what? Many of the language effects I have discussed seem subtle, to the point of perhaps being negligible.

That is one way to interpret the evidence I have discussed. Another way is to evaluate the empirical record in terms of effect sizes and their possible implications. For example, using Cohen's *d* as a yardstick, where *d* is a mean difference divided by its standard deviation, language effects on the mental accessibility of attitudes, beliefs, and so forth are often large ($d \approx .80$) (cf. Ogunnaike et al. 2010). This implies that some of language's biggest impact occurs at a deep, automatic level, influencing what is initially activated in memory (Lodge and Taber 2013).

When we turn to *reported* opinions, effect sizes across observational and experimental research often run between small ($d \approx .20$) (Pérez and Tavits 2015) and medium ($d \approx .50$) (Pérez 2011). Should analysts care about modest language effects like these? Yes, because even if they are small, they can have large ramifications. Take gender inequality (Pérez and Tavits 2015), in which language effects are reliable but lower in size ($d \approx .20$). Such effects help to illuminate why gender inequality persists in many nations despite aggregate improvements in their socioeconomic development, which is known to narrow gender gaps (Doepke et al. 2012).

What, then, should researchers do in light of small and large language effects, especially since interview language is generally omitted from statistical models of survey response? One might be tempted here to minimize, if not eliminate, the influence of language by design: for example, by pairing rigorous questionnaire translation with cognitive interviews before the full survey goes into the field. Such efforts, however, are effective only at ensuring that survey responses are comparable across different languages (i.e., measurement equivalence). That is, they are a fix to a methodological nuisance. Yet the influence of language on survey response is a theoretical proposition, one backed by scores of psychological studies and some emerging political science research. The real question, then, is how scholars can empirically account for this theoretical relationship between language and survey response.

One answer is to include interview language as a covariate in regression models of survey response. But given the challenges of this approach—for example, bloated specifications that overadjust statistical estimates—scholars might use inclusion of interview language to streamline statistical models of survey response. For example, in models of Latino opinion, native-born and citizenship status could plausibly be reinterpreted as proxies for language's distal influence, thus substituting one variable for two. Beyond simply treating language as a covariate, scholars might also consider conceptualizing language as a moderator of survey response (Baron and Kenny 1986), with interview language strengthening (weakening) the relationship between another factor (e.g., national identity) and survey response (e.g., opposition to immigration).

Nevertheless, these strategies only address the direct association between reported opinions and interview language. They do nothing about language effects further up people's cognitive stream, where the ingredients of individual opinions first come to mind (Lodge and Taber 2013; Zaller 1992). This requires looking at different outcomes, such as millisecond differences in the activation of people's mental contents. It also entails different modeling strategies, such as mediational analyses, to investigate whether the impact of language on survey response is channeled through these differences in activation (Baron and Kenny 1986; Imai et al. 2011).

In the end, however, survey researchers should care about language for theoretical, not methodological reasons. Indeed, without a more concerted effort to engage and integrate language's manifold cognitive effects into models of survey response, researchers risk misinterpreting why people report what they do in public opinion polls.

Acknowledgments

I am indebted to Cindy Kam for her incisive feedback, which enabled me to clarify key ideas in this chapter. I am also grateful for Kristin Michelitch's helpful reactions and advice on an early version of this project, as well as the editors' thoughtful and constructive feedback on the initial draft of this manuscript. I also appreciate Marc Trussler's assistance in editing this paper. Finally, I thank my young sons, Efrén III and Emiliano, for providing me with an even deeper appreciation for the power of language.

References

Abrajano, M. A., and R. M. Alvarez. 2010. *New Faces, New Voices: The Hispanic Electorate in America*. Princeton, NJ: Princeton University Press.

Almond, G. A., and S. Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton, NJ: Princeton University Press.

Baron, R. M., and D. A. Kenny. 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Bond, M. H., and K. Yang. 1982. "Ethnic Affirmation versus Cross-Cultural Accommodation: The Variable Impact of Questionnaire Language on Chinese Bilinguals from Hong Kong." *Journal of Cross-Cultural Psychology* 13: 169–185.

Boroditsky, L. 2001. "Does Language Shape Thought? Mandarin and English Speakers' Conceptions of Time." *Cognitive Psychology* 43: 1–22.

Boroditsky, L. 2003. "Linguistic Relativity." In *Encyclopedia of Cognitive Science*, edited by L. Nadel, pp. 917–921. London: Macmillan Press.

Boroditsky, L., and A. Gaby. 2010. "Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community." *Psychological Science* 21 (11): 1635–1639.

Boroditsky, L., L. A. Schmidt, and W. Phillips. 2003. "Sex, Syntax, and Semantics." In *Language in Mind: Advances in the Study of Language and Cognition*, edited by D. Gentner and S. Goldin-Meadow, pp. 61–79. Boston: MIT Press.

Brader, T., and G. E. Marcus. 2013. "Emotion and Political Psychology." In *The Oxford Handbook of Political Psychology*, edited by L. Huddy, D.O. Sears, and J.S. Levy, pp. 165–204. Oxford: Oxford University Press.

Button, K. S., J. P. A. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. J. Robinson, and M. R. Munafo. 2014. "Power Failure: Why Small Sample Size Undermines Reliability of Neuroscience." *Nature Reviews Neuroscience* 14: 365–376.

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

Chong, D., and J. N. Druckman. 2007. "Framing Theory." *Annual Review of Political Science* 10: 103–126.

Clarke, Kevin A. 2005. "The Phantom Menace: Omitted Variable Bias in Econometric Research." *Conflict Management and Peace Science* 22 (4): 341–352.

Cohen, J. 1992. "Statistical Power Analysis." *Current Directions in Psychological Science* 1 (3): 98–101.

Collins, A. M., and E. F. Loftus. 1975. "A Spreading-Activation Theory of Semantic Processing." *Psychological Review* 82: 407–428.

Cubelli, R., D. Paolieri, L. Lotto, and R. Job. 2011. "The Effect of Grammatical Gender on Object Categorization." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37 (2): 449–460.

Danziger, S., and R. Ward. 2010. "Language Changes Implicit Associations Between Ethnic Groups and Evaluation in Bilinguals." *Psychological Science* 21 (6): 799–800.

Davidov, E. 2009. "Measurement Equivalence of Nationalism and Constructive Patriotism in the ISSP: 34 Countries in a Comparative Perspective." *Political Analysis* 17 (1): 64–82.

Davidov, E., and S. Weick. 2011. "Transition to Homeownership Among Immigrant Groups and Natives in West Germany, 1984–2008." *Journal of Immigrant and Refugee Studies* 9: 393–415.

Davis, D. W. 1997. "The Direction of Race of Interviewer Effects Among African Americans: Donning the Black Mask." *American Journal of Political Science* 41 (1): 309–322.

de la Garza, R., L. DeSipio, F. Garcia, J. Garcia, and A. Falcon. 1992. *Latino Voices: Mexican, Puerto Rican, and Cuban Perspectives on American Politics*. Boulder, CO: Westview Press.

Delli Carpini, M. X., and S. Keeter. 1996. *What Americans Need to Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.

Doepke, M., M. Tertilt, and A. Voena. 2012. "The Economics and Politics of Women's Rights." *Annual Review of Economics* 4: 339–372.

Druckman, J. N., D. P. Green, J. H. Kuklinski, and A. Lupia. 2011. "Experiments: An Introduction to Core Concepts." In *Cambridge Handbook of Experimental Political Science*, edited by J. N. Druckman, D. P. Green, J. H. Kuklinski, and A. Lupia, pp. 15–26. New York: Cambridge University Press.

Dutwin, D., and M. H. Lopez. 2014. "Considerations of Survey Error in Surveys of Hispanics." *Public Opinion Quarterly* 78 (2): 392–415.

Ervin, S., and R. T. Bower. 1952. "Translation Problems in International Surveys." *Public Opinion Quarterly* 16 (4): 595–604.

Fraga, L., J. Garcia, R. Hero, M. Jones-Correa, V. Martinez-Ebers, and G. Segura. 2010. *Latino Lives in America: Making It Home*. Philadelphia: Temple University Press.

Fuhrman, O., K. McCormick, E. Chen, H. Jian, D. Shuaimei, S. Mao, and L. Boroditsky. 2011. "How Linguistic and Cultural Forces Shape Concepts of Time: English and Mandarin in 3D." *Cognitive Science* 35: 1305–1328.

Garcia, J. A. 2009. "Language of Interview: Exploring Patterns and Consequences of Changing Language During Interview." Paper presented at the Annual Meeting of the Western Political Science Association, Vancouver, BC.

Gelman, A., and G. King. 1993. "Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?" *British Journal of Political Science* 23 (4): 409–451.

Godden, D. R., and A. D. Baddeley. 1975. "Context-Dependent Memory in Two Natural Environments: On Land and Underwater." *British Journal of Psychology* 66 (3): 325–331.

Grant, H. M., L. C. Bredahl, J. Clay, J. Ferrie, J. E. Groves, T. A. McDorman, and V. J. Dark. 1998. "Context-Dependent Memory for Meaningful Material: Information for Students." *Applied Cognitive Psychology* 12: 617–623.

Harkness, J. A., F. J. R. Van de Vijver, and P. Mohler. 2003. *Cross-Cultural Survey Methods*. New York: John Wiley.

Heider, E. R. 1972. "Universals in Color Naming and Memory." *Journal of Experimental Psychology* 93 (1): 10–20.

Hochman, O., and E. Davidov. 2014. "Relations Between Second-Language Proficiency and National Identification: The Case of Immigrants in Germany." *European Sociological Review* 30 (3): 344–359.

Hong, Y., M. W. Morris, C. Chiu, and V. Benet-Martínez. 2000. "Multicultural Minds: A Dynamic Constructivist Approach to Culture and Cognition." *American Psychologist* 55 (7): 709–720.

Horn, J. L., and J. J. McArdle. 1992. "A Practical and Theoretical Guide to Measurement Invariance in Aging Research." *Experimental Aging Research* 18: 117–144.

Imai, K., L. Keele, D. Tingley, and T. Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning About Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105 (4): 765–789.

Inglehard, R., and P. Norris. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge, UK: Cambridge University Press.

Jacobson, E., H. Kumata, and J. E. Gullahorn. 1960. "Cross-Cultural Contributions to Attitude Research." *Public Opinion Quarterly* 24 (2): 205–223.

King, G., R. O. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Laitin, D. 1992. *Language Repertoires and State Construction in Africa*. Cambridge, UK: Cambridge University Press.

Lee, T. 2001. "Language-of-Interview Effects and Latino Mass Opinion." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL.

Lee, T., and E. O. Pérez. 2014. "The Persistent Connection Between Language-of-Interview and Latino Political Opinion." *Political Behavior* 36 (2): 401–425.

Levinson, S. 1996. "Frames of Reference and Molyneux's Question: Cross-Linguistic Evidence." In *Language and Space*, edited by P. Bloom, M. Peterson, L. Nadel, and M. Garrett, pp. 109–169. Cambridge: MIT Press.

Li, P., and L. Gleitman. 2002. "Turning the Tables: Language and Spatial Reasoning." *Cognition* 83 (3): 265–294.

Lien, P., M. Conway, and J. Wong. 2004. *The Politics of Asian Americans*. New York: Routledge.

Lodge, M., and C. Taber. 2013. *The Rationalizing Voter*. Cambridge, UK: Cambridge University Press.

Lucy, J., and S. Gaskins. 2001. "Grammatical Categories and the Development of Classification Preferences: A Comparative Approach." In *Language Acquisition and Conceptual Development*, edited by L. Bowermand and S. Levinson, pp. 265–294. Cambridge: Cambridge University Press.

Luskin, R. C. 1987. "Measuring Political Sophistication." *American Journal of Political Science* 31 (4): 856–899.

Marian, V., and C. M. Fausey. 2006. "Language-Dependent Memory in Bilingual Learning." *Applied Cognitive Psychology* 20: 1025–1047.

Marian, V., and M. Kaushanskaya. 2004. "Self-Construal and Emotion in Bicultural Bilinguals." *Journal of Memory and Language* 51(2): 190–201.

Marian, V., and M. Kaushanskaya. 2007. "Language Context Guides Memory Content." *Psychonomic Bulletin and Review* 14: 925–933.

Marian, V., and U. Neisser. 2000. "Language-Dependent Recall of Autobiographical Memories." *Journal of Experimental Psychology: General* 129: 361–368.

May, S. 2012. *Language and Minority Rights: Ethnicity, Nationalism, and the Politics of Language*. New York: Routledge.

McDermott, R. 2011. "Internal and External Validity." In *Cambridge Handbook of Experimental Political Science*, edited by J. N. Druckman, D. P. Green, J. H. Kuklinski, and A. Lupia, pp. 27–40. New York: Cambridge University Press.

Ogunnaike, O., Y. Dunham, and M. R. Banaji. 2010. "The Language of Implicit Preferences." *Journal of Experimental Social Psychology* 46: 999–1003.

Pérez, E. O. 2009. "Lost in Translation? Item Validity in Bilingual Political Surveys." *The Journal of Politics* 71 (4): 1530–1548.

Pérez, E. O. 2011. "The Origins and Implications of Language Effects in Multilingual Surveys: A MIMIC Approach with Application to Latino Political Attitudes." *Political Analysis* 19 (4): 434–454.

Pérez, E. O. 2013. "Implicit Attitudes: Meaning, Measurement, and Synergy with Political Science." *Politics, Groups, and Identities* 1 (2): 275–297.

Pérez, E. O. 2014. "Accented Politics: How Language Shapes Public Opinion." Paper presented at the Fall Meeting of the Symposium on the Politics of Immigration, Race, and Ethnicity (SPIRE) at the University of Pennsylvania.

Pérez, E. O., and M. Tavits. 2015. "His and Hers: How Language Shapes Public Attitudes Toward Gender Equality." Paper presented at the Annual Meeting of the Midwest Political Science Association.

Pérez, E. O., and M. Tavits. n.d. "Today Is Tomorrow: How Language Shifts People's Time Perspective and Why It Matters for Politics." Unpublished manuscript, Vanderbilt University.

Pew Research Center. 2014. "Attitudes about Aging: A Global Perspective."

Piston, Specer. 2010. "How Explicit Racial Prejudice Hurt Obama in the 2008 Election." *Political Behavior* 32 (4): 431–451.

Portes, A., and R. G. Rumbaut. 2006. *Immigrant America: A Portrait*. Berkeley: University of California Press.

Ralston, D. A., M. K. Cunniff, and D. J. Gustafson. 1995. "Cultural Accommodation: The Effect of Language on the Responses of Bilingual Hong Kong Chinese Managers." *Journal of Cross-Cultural Psychology* 26: 714–727.

Rosch, E. 1975. "Cognitive Representations of Semantic Categories." *Journal of Experimental Psychology: General* 104 (3): 192–233.

Ross, M., W. Q. E. Xun, and A. E. Wilson. 2002. "Language and the Bicultural Self." *Personality and Social Psychology Bulletin* 28 (8): 1040–1050.

Ryan, C. 2013. *Language Use in the United States: 2011*. Washington, D.C. U.S. Census Bureau.

Schmid, C. L. 2001. *The Politics of Language: Conflict, Identity, and Cultural Pluralism in Comparative Perspective*. Oxford: Oxford University Press.

Sears, D. O. 1986. "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature." *Journal of Personality and Social Psychology* 51 (3): 515–530.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.

Slobin, D. 1996. "From 'Thought and Language' to "Thinking for Speaking." In *Rethinking Linguistic Relativity*, edited by J. Gumperz and S. Levinson, pp. 70–96. Cambridge, UK: Cambridge University Press.

Smith, T. W. 1987. "That Which We Call Welfare by Any Other Name Would Smell Sweeter: An Analysis of the Impact of Question Wording on Response Patterns." *Public Opinion Quarterly* 51 (1): 75–83.

Stegmueller, D. 2011. "Apples and Oranges? The Problem of Equivalence in Comparative Research." *Political Analysis* 19: 471–487.

Stern, E. 1948. "The Universe, Translation, and Timing." *Public Opinion Quarterly* 12: 711–715.

Stimson, J. A. 2004. *Tides of Consent: How Public Opinion Shapes American Politics*. Cambridge, UK: Cambridge University Press.

Swoyer, C. 2014. "Relativism." In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. .

Tillie, J., M. Koomen, A. van Heelsum, and A. Damstra. 2012. "EURISLAM—Final Integrated Report." European Commission: Community Research and Development Information Service (CORDIS). .

Tourangeau, R., L. J. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.

Trafimow, D., and M. D. Smith. 1998. "An Extension of the "Two-Baskets" Theory to Native Americans." *European Journal of Social Psychology* 28: 1015–1019.

Trafimow, D., H. C. Triandis, and S. G. Goto. 1991. "Some Tests of the Distinction Between the Private Self and the Collective Self." *Journal of Personality and Social Psychology* 60: 649–655.

Triandis, H. C. 1989. "The Self and Social Behavior in Different Cultural Contexts." *Psychological Review* 93: 506–520.

Triandis, H. C., and E. M. Suh. 2002. "Cultural Influences on Personality." *Annual Review of Psychology* 53: 133–160.

Tulving, E., and D. Thomson. 1973. "Encoding Specificity and Retrieval Processes in Episodic Memory." *Psychological Review* 80 (5): 352–373.

Vigliocco, G., D. P. Vinson, F. Paganelli, and K. Dworzynski. 2005. "Grammatical Gender Effects on Cognition: Implications for Language Learning and Language Use." *Journal of Experimental Psychology: General* 134 (4): 501–520.

Villar, A., and J. A. Krosnick. 2011. "Global Warming vs. Climate Change, Taxes vs. Prices: Does Word Choice Matter?" *Climactic Change* 105: 1–12.

Welch, S., J. Comer, and M. Steinman. 1973. "Interviewing in a Mexican-American Community: An Investigation of Some Potential Sources of Response Bias." *Public Opinion Quarterly* 37(1): 115–126.

Whorf, B. L. 1956. *Language, Thought, and Reality*. Cambridge, MA: MIT Press.

Wong, J., S. K. Ramakrishnan, T. Lee, and J. Junn. 2011. *Asian American Political Participation: Emerging Constituents and Their Political Identities*. New York, NY: Russell Sage Foundation.

Zaller, J. 1992. *The Nature and Origins of Mass Opinion*. Cambridge, UK: Cambridge University Press.

Notes:

([1]) On the relevance of language for politics beyond the United States, see Laitin (1992); May (2012); Schmid (2001).

([2]) Many studies of measurement equivalence have a cross-national focus, since comparisons of countries on latent traits is an area in which a lack of equivalence is likely (Davidov 2009; Stegmueller 2011). But in this research, language is only one of many possible reasons for a lack of equivalence. Still, the logic and criteria guiding cross-national analyses of measurement equivalence also guide similar tests in cross-language settings (Pérez 2011).

([3]) To diagnose measurement equivalence (e.g., multigroup confirmatory factor analysis), researchers often need multiple measures of a trait. Yet such data are scarce, since scholars must weigh the inclusion of multiple items for single traits against space limitations, respondent fatigue, and so forth. Further, even when such data exist, analyses of equivalence only reveal whether items meet this criterion (Davidov 2009). Some methods can statistically correct a lack of equivalence (Stegmueller 2011), but these do not fully clarify what language features yield nonequivalence.

([4]) This entails formally verifying measurement equivalence across languages. Scholars can also work toward measurement equivalence in the design stage by appraising the quality of their questionnaire via pretesting, such as cognitive interviews with a small set of respondents, which can identify translation problems (Harkness et al. 2003).

([5]) This is not to mention the possible complex interactions between these groups and language (e.g., as in a phenotypically light Latino interviewing a phenotypically dark Latino in English versus Spanish).

Efren Perez

Vanderbilt University, Department of Political Science