Multilingualism and European Identification.

Author: Juan Díez Medrano

Sociological Inquiry. DOI: 10.1111/soin.12211.
Abstract:
The literature on ethnicity and nationalism notes the centrality of language in group formation. Developments connected to globalization have simultaneously propitiated learning more than one language and the emergence of post-national and pan-ethnic identifications. This article brings the literature on ethnicity and nationalism to bear on the analysis of these changes to assess the causal relationship between multilingualism (ML) and European identification (EI). Sociology and social psychology provide solid theoretical arguments for why they should be related. The article tests this relationship with *Eurobarometer* data from 2010. I show that ML has a causal impact on EI and that interaction helps mediate this relationship. The effect of language on EI is modest, however, and interaction does not appear to be the main mechanism mediating it. The article concludes with speculation as to the role of low relative cross-national mobility in Europe as the main explanation for some of the findings.

---

Does speaking various languages contribute to the development of hyphenated, pan-ethnic, supranational, or cosmopolitan identifications? If so, is this contribution primarily channeled through interpersonal interaction and communication with people from other ethnic or national origins? This article focuses on supranational identification in the European Union (i.e. European identification) to address these questions.

Establishing and explaining the role of multilingualism in the development of identifications that merge or combine ethnic identifications or transcend the nation is theoretically and policy relevant. Theoretically, because it fills important gaps in the literature on the relationship between language and identity and on the social and cultural transformations associated to globalization. Although this literature shows that
globalization both provides incentives to learning more than one language and contributes
to the emergence of new identifications, it rarely explores the links between these two
simultaneous developments. Also, although it has established strong links between
language and identification, it does not pay as much attention to the identity implications of
speaking several languages, especially for autochthonous populations.

In the context of the European Union, while a small number of publications include
the ability to speak more than one language as an independent variable in statistical models
that predict European identification, its role is under-theorized and researchers have not
systematically explored the causal structure underlying the observed association between
the two. The prevailing assumption is that fluency in foreign languages facilitates
interpersonal contact with nationals from other European countries, which in turn instils a
greater sense of being European (e.g. Kuhn, 2015; Recchi, 2014; Fligstein, 2008; Mau,
2007). This emphasis on interaction reflects recent sociological interest in practices and
experience in European integration studies (e.g. McNamara, 2015; Favell and Guiraudon,
2011). One of this article’s main goals is to test the validity of this theoretical proposition
and interpret the statistical findings in light of this and complementary theoretical
perspectives that have been applied to the study of ethnic identifications in the United
States and to the analysis of cultural assimilation among immigrants.

The questions above are politically and policy relevant because the prevalence of
multilingualism is sensitive to government regulation and public resources invested in
foreign language instruction. The examination of the impact of multilingualism on the
strength of national attachments and on the emergence of hyphenated, pan-ethnic,
supranational, and cosmopolitan identifications can thus be of interest to those who for
political (e.g. European Union leaders) or normative reasons (e.g. Beck’s *Cosmopolitan*
Vision, 2006) advocate and promote such emergence (on the European Union’s promotion of multilingualism, see Wodak, 2011; Zappettini, 2014). The results of this analysis may help shape their views on the impact of public investment in foreign language instruction (see Calligaro, 2013 for analysis of how the European Union’s cultural policy has been partly guided by the desire to promote European identification in the population). At the same time, the analysis of the effect of multilingualism on identification can put into perspective emphasis by a growing number of contemporary political organizations and intellectuals (especially those located on the far right) on the threat to national identity posed by the spread of foreign languages. Alleged lack of linguistic assimilation among immigrants, the state’s failure to protect minority languages—as in the case of regional movements—, the colonization of national languages by English, and the priority attached to English-language publications in academic evaluations, are some of the grievances that motivate these actors’ mobilization in defense of the national language.

I use the Eurobarometer study 73.3 to show that knowing various languages is not just statistically correlated with European identification, as previous research has shown, but in fact causally contributes to making its citizens more aware that they are European. I reach this conclusion after estimating statistical models that tackle measurement and specification problems, such as the possibility that European identification impacts on people’s motivation to learning languages. Then, I enter into yet unexplored territory to empirically evaluate the relative significance of interpersonal interaction in mediating this relationship. The statistical results challenge prevailing assumptions in the contemporary study of European identification by showing that interpersonal interaction is at best only part of the explanation for the relationship between multilingualism and European identification. It is thus conceivable that other mechanisms
highlighted in the literature on language and identity, such as perceptions of cultural
distance from primary speakers of languages different from one’s mother tongue and
people’s desire to symbolically differentiate themselves from other members in their
community or satisfy their need to be from ‘somewhere’ (Waters, 1990) also connect the
two variables. Data constraints, however, preclude an empirical test of the role that these
alternative mechanisms play.

A theoretical and empirical sections structure the article. The theoretical section
briefly situates the study of the association between multilingualism and hyphenated, pan-
ethnic, supranational, and cosmopolitan identifications in the literature on language,
identity, and globalization. Then, it synthesizes two theoretical approaches, one stressing
interaction/communication, the other stressing socio-psychological processes, which
theoretically ground the expectation of a causal relationship between multilingualism and
European identification. The empirical section first describes the EB73.3 data and the
variables included in the statistical analysis. Then it examines the relationship between
multilingualism and European identification and evaluates the mediating role of interaction.

Language and Identification in the Age of Globalization

Numerous authors (e.g. Gellner, 1983; Anderson, 1983; Hobsbawm, 1992; Brubaker, 2013)
have noted that the centrality of communication for economic and political organization in
modern society turned language into a major cultural marker of ethnic and national identity,
even though it is not necessarily the main or only one (e.g. López and Espiritu, 1990;
Cornell, 1988; Wimmer, 2002)

The onset of the second wave of globalization has transformed society’s linguistic
requirements and multiplied opportunities to learn languages (e.g. Gerhards, 2012; Luo and
Shenkar, 2006). In a world in which transnational corporations operate in different
countries, where international exchange increases exponentially, where millions of people are able and willing to move for short or long periods of time, and where long-distance communication across space is both easy and cheap, the demand for and prevalence of people who are fluent in various languages have increased dramatically.

Because of the centrality of language in the development of ethnic and national identifications noted above one expects that many among the rising number of people who are fluent in various languages face serious dilemmas when choosing their identity. It is conceivable, in fact, that many solve this dilemma by distancing themselves from the primary identifications into which they were socialized and adopting new ones.

The thriving literature on new ethnic or post-national identifications (Delanty, 2012; Jung, 2008; Beck, 2006; for a general overview of this literature see Vertovec and Cohen, 2002), however, has paid relatively little theoretical and empirical attention to the role of multilingualism. One exception to this is research on hyphenated identifications and on pan-ethnicity in the United States, which has explored the ways in which language can contribute to the formation of new ethnic identities from already existing ones (e.g. Waters, 1990; Espiritu, 1992). Espiritu, for instance, discusses the role of having a common language of communication, English, in unifying various Asian nationalities into a common Asian identity in the United States. Another exception is recent research on Europe that focuses on transnational practices. This research, however, rarely situates language at the center of the investigation (for an exception, see Fuss et al., 2004) and does not elaborate on the causal mechanisms that mediate between multilingualism and European identification. More theoretical reflection and empirical research are thus needed on the impact of language on people’s hyphenated, pan-ethnic, supranational, and cosmopolitan identifications.
Multilingualism and Identity in the European Union

This article focuses on European identification, the archetypical example of a supranational identification (e.g. Kaina, 2013; Mols and Weber, 2013; Cram, 2012; Herrmann et al. 2004; Bruter, 2004; Duchesne and Frognier, 1995). Scholars have noted that polities play a very important role in constructing and delimiting the boundaries of the categories with which people identify (e.g. Kim and White, 2010; Davis, 2001; Breuilly, 1994). This is relevant because the word “Europe” refers to multiple realities, geographic, economic, political, and cultural, whose precise boundaries are fuzzy and non-overlapping (Eder, 2006). The 1993 Treaty of European Union’s institutionalization of European citizenship gave a major boost to the salience of “Europe” as a category of identification. Citizenship in a European Union member state automatically confers the status of European citizen. The institutionalization of the category European has increased its salience as a potential object of identification, situated above national and ethnic identifications that are contained within it.

Authors distinguish between a cultural and a civic dimension of European identification (Cram, 2012; Bruter, 2004). Bruter, in particular, notes that the ‘cultural’ component of identity represents, by and large, the sense of belonging of an individual citizen towards a particular group. The ‘civic’ component, on the other hand, has to do with the citizens’ identification with a given political structure (Bruter, 2004, p. 26). I focus on the cultural dimension of European identification, and more specifically on the importance that citizens attach to being European, in order to better connect this investigation to the sociological study of the significance of language for group formation in general (e.g. see Brubaker, 2013).

Just above fifty-percent of European Union citizens express a sense of identification as European, but less than ten percent describe themselves as “only” European. The rest of
the population retains an exclusive national or regional identification (Risse, 2011; Fligstein, 2008). Explaining European identification is in some ways similar to explaining hyphenated (e.g. German-American) and panethnic identifications (e.g. Latino), as in the US, and national and regional identifications in plurinational states (e.g. Scottish and British) (Okamoto and Mora, 2014). This family resemblance is especially strong when the focus is on a group’s adoption of an identification that subsumes one’s original identification into a larger one (e.g. from German immigrant into German-American or, simply, American; from Vietnamese or Vietnamese-American into Asian-American or just Asian; from Quebecois into Canadian or French-Canadian). Contributing to the explanation of European identification can thus illuminate aspects connected to related topics, covered by distinct literatures. The discussion above suggests that speaking one or more languages different from one’s mother tongue should trigger processes that lead individuals to develop new identifications. The main challenge for researchers is to ground this expectation theoretically, empirically disentangle this effect from the reciprocal influence that European identification can have on the decision to learn new languages (Rivers, 2011; Hochman and Davidov, 2014), and test for the mechanisms that mediate the causal relationship, if there is one.

The literature on the relationship between language and identification has mainly focused on nationalist and regionalist movements and on the assimilation of migrants to host societies and the concomitant adoption of the host country’s identification. The section below discusses theories or arguments borrowed from these literatures to explore possible connections between multilingualism and supranational identification among native populations. One can distinguish a strand that emphasizes the mediating role of
interaction and another one that emphasizes the mediating role of socio-psychological processes—cognitive and affective.

Language, Interaction, and European Identification

Authors who have explored the assimilation of immigrants to host societies have often drawn on the contact hypothesis (Allport, 1954; Pettigrew and Tropp, 2006). The contact hypothesis states that positive interaction and cooperation between members of different groups leads to reduced prejudice, greater empathy, and mutual positive feelings. The literature has inferred from this hypothesis that learning the host population’s language facilitates the immigrants’ interaction with members of this population and thus makes it more likely that they eventually adopt the host population’s category of identification, irrespective of how attached they remain to their group of origin (e.g. Amit and Bar-Lev, 2015). One can apply the contact hypothesis to the analysis of the effect of multilingualism on European identification. Fluency in foreign languages increases opportunities for interaction with nationals from other European countries, which over the course of time increases positive feelings toward them and a growing sense of being European.

Deutsch provides a more elaborate and general account of the role of language in the development of identifications. This account focuses on the bottom-up emergence of new collective identifications that results when previously unacquainted members of different groups sever their ties with their communities of origin and start to interact with one another, as was the case during the period of fast urbanization in 19th century Europe. In Nationalism and Social Communication (1953), Deutsch places communication (and the theory of communication that was dominant in the 1950s) at the center of the explanation of the emergence of peoples and nations and, by implication, of a person’s sense of belonging. Language enters his explanation by making sustained interaction across domains and
situations possible, which eventually produces “effective communication”, a quasi-perfect understanding between people over the widest possible range of issues (Pp. 96-97, p. 98)

Like many theorists of nationalism who have emphasized language, Deutsch does not address the topic of multilingualism. His argument implies, however, and this is how specialists on European identification see it (e.g. Kuhn, 2015; Fligstein, 2008; Mau, 2007), that knowing another group’s language facilitates interaction with members of this other group, which when sustained over a long period of time, should lead to the emergence of a shared culture and identity among the parties to interaction, without necessarily undermining one’s primary attachments. In the European Union context, it seems natural that individuals who participate in a shared culture with nationals from other European Union member states would come to define themselves as European, for this is the most salient and institutionalized identity situated above the national level (on people’s preference for nested than cross-cutting categories of identification, see Chai, 2005).

In sum, both the contact hypothesis and Deutsch’s communication theory discussed above state that language connects individuals through interaction, which in turn generates new identifications.

*Cultural Competence and Identity*

While it makes sense to think that interaction plays a role in channeling the effect of speaking several languages on European identification, other theoretical and analytical perspectives imply that socio-psychological processes unrelated to interaction also lead European multilinguals to identify more as European than do those who only speak their mother tongue, and may even play a more important role. These perspectives have been overlooked in contemporary accounts of European identification.
A distinctive aspect of European identification that sets it apart from most ethnic or national identifications, is that it is so far devoid of emotional content (e.g. Duchesne et al. 2013; Guibernau, 2011). Like hyphenated ethnic identifications among whites in the United States (see Waters, 1990; see also Cornell and Hartmann, 1998), European identification can thus be characterized as “thin” and largely a matter of choice. Just as having German roots has little impact on the lives of American citizens, the fact of being European still plays a relatively small role in determining the European citizens’ lives. The most tangible benefit they can perceive is that they do not need passports when travelling to European Union member states who have signed to the Schengen Agreement. In many ways, European Union citizens are in a similar situation to that of whites in the Unites States (see Alba, 1990; Waters, 1990; Gans, 1979), for whom identification with an ethnic or national category is largely a matter of choice, a choice made for symbolic reasons, to fulfil the need to be from “somewhere” and, more generally, to stand out as an individual (Waters, 1990). One could thus say that self-categorization as European is also a choice, and just as white Americans make ethnic choices based on idiosyncratic information that connects them to a particular group, one can surmise that European Union citizens attach more importance to their European identity when they possess traits that one easily recognizes as defining of what Europe is. Thus, while the choice is open to all European Union citizens, it is conceivable that multilinguals, because of language’s centrality in the definition of who people are (Brubaker, 2013) and because language diversity is one of the European Union’s most salient cultural features, are more aware of being European and feel more inclined and entitled to call themselves Europeans than are and do speakers of only one language.
Other theories reach beyond the symbolic and call upon more specific socio-psychological processes to justify the impact that language has on identification. Some emphasize closeness to others who speak the same language. When applied to European identification, one could argue that speaking several languages brings people closer to the country and the people who speak those languages, which makes the institutionalized category of identification “European” more salient to them. Others emphasize that speaking several languages provides access to knowledge of other cultures, a feeling of multicultural competence that enhances the multilinguals’ self-esteem and induces them to identify as part of the same community as those who speak the different languages that they command (e.g. Gaudet and Clément, 2009; Noels et al., 1996; Kim, 1988).

The most elaborate theoretical discussions of the socio-psychological impact of speaking several languages on individuals’ identification draw on Social Identity theory and on the Bicultural Identity Integration (BII) paradigm to stress the role of the reduced perceived cultural distance that comes with knowing a group’s language (see also Gordon, 1964; Taft, 1957). The former underlines that holding traits or characteristics that one considers as markers of group identity lead to perceptions of cultural similarity to the members of the group being considered and feelings of membership in this group (e.g. Turner, 1987). The latter elaborates this hypothesis further and with a specific focus on the role of language. Competence in the dominant languages of the cultures in which bicultural and multicultural individuals live leads to perceiving less distance between cultures and to integrating or blending identifications (Chen et al., 2008; Benet-Martínez and Haritatos, 2005, p. 1033; Haritatos and Benet-Martínez, 2002).

The socio-psychological approaches above have been mainly concerned with immigrant populations. This means that in order to apply them to the study of European
identification in the European Union, one must translate and reformulate their questions and predictions. Instead of asking whether and to what extent immigrants integrate the culture of origin and the culture of destination or what makes citizens with immigrant ancestry retain this immigrant ancestry as part of a hyphenated identification, one must ask “how central is European identification to nationals of states located in Europe? What leads nationals of these European countries to supplement their national identification with a strong sense of being European?” The implications that follow from Social Identity Theory and the BII paradigm is that European citizens integrate their national and European identity, the greater their feeling of competence in European languages that are spoken in countries other than their own. Not so much because of enhanced self-esteem, as one theory above emphasizes, but because competence in foreign languages reduces the cultural distance that people perceive between Europeans of different nationalities. In qualitative studies, Díez Medrano (2003) and Bruter (2003) have in fact noted a strong connection between people’s perception of a greater or lesser cultural homogeneity across Europe and the extent to which they feel European. What the BII paradigm in particular does is to highlight the role of language competence in explaining this perception of cultural homogeneity.

The discussion above provides different reasons, anchored in socio-psychological processes, to expect a relationship between multilingualism and European identification: people’s need to stand-out from the crowd, people’s greater affinity to other cultures and peoples, people’s sense of competence in other cultures and the enhanced self-esteem that follows from this, and people’s perception that individuals and countries from other countries are not so different in cultural terms. Because of this, the socio-psychological tradition, like the sociological one, predicts that the more languages Europeans speak, the
greater their sense of being European. In contrast to what the contact hypothesis and Deutsch’s theory of group formation predict, however, the socio-psychological tradition discussed in this section stresses that the effect of knowledge of foreign languages on a person’s sense of belonging is not necessarily and not primarily mediated through interaction; symbolic factors and competence in cultures other than one’s own, and the resulting reduction in perceived distance between these cultures, also play a role.

In sum, various sociological and socio-psychological approaches lead to the theoretical proposition that *speaking more than one language causally contributes to European identification among individuals.* In the empirical analysis below I explore the validity of this proposition through a test of the following hypothesis:

**H1: The more languages Europeans speak, the greater their sense of being European.**

Then, I examine the role of interpersonal interaction in mediating the association between multilingualism and European identification. The sociological approach puts interaction at the center of the explanation whereas other approaches are more agnostic as to the primary role of interaction. My examination of the role of interpersonal interaction must by necessity be inferential, for the data used in this article do not provide information on actual interpersonal interaction. Also, the cross-sectional and non-experimental character of the data makes statements concerning causal chains such as this problematic. To circumvent these problems, I take a conservative approach. I focus on not one but three hypotheses, which one can test with the dataset below and are related and logically consistent with the sociological approach’s theoretical proposition that *the main causal link between the number of languages one speaks and European identification is the frequency of interaction with nationals from other European countries.*
Hypothesis 2 simply reformulates the theoretical proposition, using plausible proxies for actual interpersonal interaction with other Europeans.

**H2:** *The main reason why people who speak more languages identify more as European than do those who speak fewer languages is that speaking several languages allows them to spend holidays, work, and live abroad more frequently.*

Then, hypotheses 3 and 4 are corollaries that should be empirically true if the main theoretical proposition is also true.

**H3:** *The greater the number of Europeans who speak an individual’s national language (i.e. the greater its communicative value), the stronger this individual’s European identification.*

The rationale behind this hypothesis is that the greater the communicative value of an individual’s national language is, the more frequent the speaker of this national language’s interaction with other Europeans and, because of this more frequent interaction, the stronger his or her European identification.

**H4:** *The causal impact of knowing a particular foreign language on an individual’s European identification is stronger, the greater this language’s prevalence in the European Union.*

The rationale behind this hypothesis is that knowing foreign languages that are more widely spoken allows for more interaction with Europeans than does knowing foreign languages that are less widely spoken. This, in turn, leads those who know a widely spoken foreign language to identify more as European than do those who know less widely spoken foreign languages.

Data and Methods
I use data from the *Eurobarometer* study 73.3 (2010). *Eurobarometer* is a yearly cross-sectional representative survey of EU citizens 15 years old and older. The study 73.3 asks questions on the languages that respondents speak, on transnational practices, and on European identification. To test the hypotheses above, I estimate OLS, Hierarchical, and Instrumental Variable models based on the information contained in this study.

**The Effect of Fluency in Foreign Languages on European Identification**

In the analysis that follows I only focus on the subsample of respondents born in the country of residence. Non-random selection processes influence whether a person migrates and stays in another country, which are correlated with both the number of languages that people speak (thereafter ML, short for MultiLingualism) and European identification (thereafter EI).

**Variables**

The *Eurobarometer* questionnaire asks respondents about their mother tongue and then about all the languages that they speak well enough to have a conversation. Until now, this has been the main source of comparative data on ML across Europe (e.g. De Swaan, 2002; Gerhards, 2012), although self-reports of ML are often inaccurate and include systematic bias (Edele et al., 2015). I use the total number of languages that respondents speak as the main independent variable. Then, when testing the role of interaction in mediating the relationship between ML and EI, I include other language variables in the statistical models. I use a measure of the national languages’ communicative value, the Q-Index, to test H3. De Swaan calculates this measure on the basis of self-reports taken from the *Eurobarometer* study. Its values correspond to the product of the prevalence and the centrality of the different national languages in the European Union (De Swaan, 2002)\(^1\). Also, when testing for the role that knowing widely used foreign languages plays in
connecting people from different nationalities (H4), I use two additional language variables, which distinguish between respondents who speak English as a foreign language, respondents who speak foreign languages other than English, and respondents who only speak their mother tongue.

_Eurobarometer_ allows for the analysis of the cultural dimension of European identification and thus a test of the sociological and socio-psychological hypotheses above. Its questionnaire includes the following query:

“Thinking about the fact that you are European, how important is being European for you personally?”

Respondents had a choice between the answers “A lot”, “Somewhat”, “Not much”, and “Not at all”.

A positive aspect of the EB 73.3 question when compared to alternative versions included in other surveys (for a critique of these measures, see Citrin and Sears, 2009, p. 149) is that it directly speaks to the meaning and significance that the category of belonging “European” has for an individual. Like the measure of ML, however, the measure of EI also suffers from measurement error. This is in part because the salience of Europe as a category of identification pales in comparison with that of smaller geopolitical units and social categories (Risse, 2011). Another measurement problem is that people mean quite different things when they report on their level of EI (Cram, 2012; Bruter, 2004; Grundey and Jamieson, 2007).

All but one of the statistical models below control for the respondents’ country of residence. In addition to this, the models include three individual-level socio-demographic control variables: Gender, Age, and Level of education. Gender is a dichotomous variable, with the value 1 representing male respondents (see Mau, 2007; Kuhn, 2015). Age
represents how old respondents are (see Jung, 2008; Down and Wilson, 2013). Education measures how old respondents were when they completed full-time education (see Inglehart, 1970; Kuhn, 2015).

The *Eurobarometer* 73.3 collected information on people’s transnational ties and activities, through a long battery of questions to which respondents simply had to answer “yes” or “no”. The statistical analysis includes three variables based on these questions, in order to test whether transnational experiences and the expected interaction with other Europeans that they make possible mediate the relationship between ML and EI (H2). The variables measure whether respondents have resided, worked, or spent holidays abroad. We know that most of these experiences indeed take place within Europe (e.g. Recchi, 2015). One problem with the three measures is that they do not convey information on the amount and quality of interaction with other Europeans that resulted from working, living, or spending holidays abroad.

To test H3, I use the Q-Index described above. Its inclusion in a country fixed-effects model is problematic, however, due to the fact that it is collinear with the country of residence dummy variables. To overcome this problem, I test H3 by estimating a random-slopes two-level hierarchical model instead of a country fixed-effects model and by replacing the country of residence dichotomous variables with several country-level variables that serve as control variables in the estimation of the relationship between the Q-Index and EI. These control variables are GDP per capita (Kuhn, 2015), Protestantism (Leustean, 2013; Boomgaardon and Freire, 2009; Nelsen et al., 2001), and GDP Growth.

Although the hypothesis that ML impacts on EI makes intuitive sense, to empirically ascertain its validity is far more complicated because of systematic measurement error and possible omitted variable bias. More specifically, just as ML can
increase EI, EI can motivate people to learn new languages or to exaggerate their ability to speak different languages. Self-verification theory, for instance, notes that people adapt their behavior to their values and convictions (Swann and Read, 1981). It is thus conceivable that individuals with a strong attachment to their nation would resist learning languages other than their mother tongue, as Rivers has shown for Japan (Rivers, 2011). These problems, often encapsulated under the term “endogeneity”, lead to biased and inconsistent estimates of causal effects (Cameron and Trivedi, 2005). To obtain a consistent estimate of the causal effect of ML, I complement standard OLS regression with single-equation instrumental regression (IV). The instrument that I use in this article is one of the items included in the battery of questions on people’s transnational experiences and connections mentioned above. The item refers to whether respondents have close relatives (brothers, sisters, children, parents) who live abroad. Whereas relatives who live abroad can been seen as providing opportunities to learning other languages (e.g. through trips abroad, where respondents are exposed to these other languages), the fact that they share the respondents’ national or ethnic origins makes it unlikely that they influence the latter’s sense of being European.

Analysis

Table 1 displays descriptive statistics for all the variables in the models. Then, Table 2 presents the results of two statistical models that measure the relationship between ML and EI. Column 1 provides an overall picture of the association between ML and EI, when holding other variables constant, including the respondents’ country of residence. As expected, ML is positively related to EI. To speak one additional language is associated with a 0.06 units increase in the indicator of European identification. This is equivalent to an increase of 6% in the inter-quartile range for the dependent variable, which goes from 1
to 2 in a 0-3 scale. I considered the possibility of non-linearity in the effect of ML on EI. The test that compares the difference in fit between the linear and non-linear formulations, however, is not statistically significant. The overall fit of the model reported in Column 1 is small (R-square = 11.1%) but in line with other findings (Kuhn, 2015). Country of Residence alone accounts for 8.2% of the variance (not shown).

[Tables 1 and 2 about here]

Column 2 displays the IV coefficients for the same model as in Column 1. The coefficient for the effect of ML is considerably bigger than the one obtained with the actual measure of ML. A value of 0.32 for ML’s coefficient means that for every additional language that people speak, the value for the importance that they attach to being European increases by a third of this variable’s inter-quartile range.

The credibility of the findings obtained through IV models hinge to a large extent on the quality of the instruments. What follows adheres to standard recommendations about how to check for this quality. First of all, the correlation index for the association between ML and Has Relatives Abroad is 0.11, whereas the corresponding index for the association between Has Relatives Abroad and EI is trivial, at 0.01. Post-estimation statistics also reveal 1) that the unique contribution of the excluded instrument (Has Relatives Abroad) to the explanation of ML is small (0.19) but statistically significant, 2) that the F-value (97.3) for the model that predicts ML with all the exogenous variables (individual-level variables, the country dichotomous variables, and the excluded instrument “Has Relatives Abroad”) is much greater than 10 and statistically significant (Stock and Yogo, 2005), and 3) the test of the null hypothesis that ML is exogenous instead of endogenous is statistically significant at 0.05, two-tailed. The post-estimation tests thus
endorse the use of the variable “Has Relatives Abroad” as an instrument for ML (see Table 2).

In all, the results above are consistent with the theoretical expectations derived from the interaction and the socio-psychological approaches to identification. They strongly support the hypothesis that ML has an independent and substantial causal impact on EI.

Interaction-Communication and the impact of ML on EI

While the literature has explored the relationship between knowing languages and identification as European, it has never tested propositions concerning the variables that mediate and therefore explain this relationship. The information contained in the Eurobarometer study 73.3 advances our understanding of the mechanisms that mediate the association between ML and EI by allowing for several indirect tests of the relative significance of interaction.

Table 3 displays four different models that test hypotheses H2, H3, and H4. A comparison of the coefficients displayed in Table 2, column 1, and those in Table 3, column 1, helps evaluate the extent to which transnational experience mediates the relationship between ML and EI. It shows that the coefficient for ML drops from 0.06 to 0.04 (or 33%), but remains statistically significant. It also shows that the effects of having worked abroad (0.05) and spending holidays abroad (0.09) are in the predicted direction and statistically significant. A plausible, but not the only, interpretation of these findings, is that the effect of ML on EI is partly mediated through interaction with other Europeans. Knowing languages facilitates mobility abroad, which promotes contact with other Europeans, and eventually increases a person’s sense of being European. The fact, however, that the coefficient for the direct effect of ML on EI is still positive and
statistically significant is consistent with the possibility that factors other than interaction also intervene in the relationship.

Table 3, Column 2, displays the coefficient estimates from a random-slopes model that explores the relationship between one’s national language’s communicative value and European identification. The statistical results contradict H3, by showing that the sign of the relationship between the Q-Index and EI is opposite to what one would expect. In addition to this, column 2 shows that the effect of ML on EI is still positive and statistically significant after controlling for the Q-index and the other individual and country-level variables in the model.

Emphasis on the role of interaction in mediating the impact of language on identification also leads to the expectation that the effect of ML on EI is stronger when one speaks a foreign language with high communicative value than when one speaks a language with low communicative value. I compare the effect of knowing English and other languages to that of knowing languages other than English. To differentiate this test from the previous one, which focused on the communicative value of one’s national language, I exclude respondents from English speaking countries. The statistical results reported in columns 3 and 4 contradict theoretical expectations, by showing that there is no difference between speaking English and speaking other languages.

The results of the tests concerning the impact of the national and foreign languages’ communicative value on EI contradict the expectation that interaction and communication mediate the effect of languages on EI. One possible explanation for these findings is that nationals of countries whose language has high communicative value and individuals who speak English as a foreign language have less transnational experience than have nationals
of countries whose language has lower communicative value and individuals who speak foreign languages other than English. In other words, it is conceivable that the potential for greater interaction with foreign Europeans entailed by speaking widely spoken national and foreign languages is not actualized. The statistical results displayed in Table 4 discard this possible interpretation. Table 4 reports regression coefficients for different models that estimate the effects of the national language’s communicative value and of speaking English as a foreign language versus speaking foreign languages other than English on the probability of having worked, lived, or spent holidays abroad. Columns 1 to 3 show that the greater a national language’s communicative value, the greater the probability that individuals have done any of those things, whereas Columns 4 to 6 show that whether individuals in non-English speaking countries speak English or other foreign languages has no effect on the probability that they have worked, lived, or spent holidays in other countries. Meanwhile, all columns confirm that the more languages individuals speak, the more often they have engaged in those transnational activities.

Discrimination and Conclusion

Sociological and socio-psychological theories of identification invite the prediction of a causal relationship between multilingualism and European identification. This article confirms this expectation (H1) and suggests that ML’s effect is much bigger than previously thought. In addition to this, the article moves us beyond the state of the art, by examining the role of interaction in mediating this causal relationship. Some statistical findings support the hypothesis that interaction is a major mechanism intervening in the relationship while others suggest that it is neither the only nor the dominant one. Speaking for the role of interaction is the finding that the effect of ML on EI is 33% smaller when
one controls for Work, Residence, and Holidays Abroad (H2). This finding is important because it suggests that the Europeans’ relatively sedentary habits (see Recchi, 2015) may be short-circuiting ML’s potential for generating EI.

Whereas the finding above suggests that interaction mediates the effect of ML on EI, the lack of a relationship between the communicative value of the national and foreign languages that a person speaks and the importance that he or she attaches to being European (H3 and H4) questions the primacy of interaction as the intervening causal mechanism between ML and EI. This opens the possibility for another interpretation of the statistical findings in this article, which is that in the European context, where there are many more “multilinguals” than “transnationals”, socio-psychological factors also matter, and matter more perhaps in the explanation of the effect of ML on EI than does interaction. These socio-psychological factors could account for a large share of the effect of ML on EI after controlling for measures of transnational experience and for why the national and foreign languages’ communicative value does not impact on EI as the interaction perspective would predict. More research is of course needed to support this alternative interpretation.

The article expands our knowledge about the impact of language on identity in the context of globalization. The statistical findings suggest that learning languages widens people’s life and mental horizons and facilitates the development of cosmopolitan identifications, by facilitating interaction certainly, but quite likely also for symbolic reasons and by reducing perceptions of inter-cultural distance and increasing people’s appreciation of other people and countries. Thus, while experience, practices and sociation matter, cognition and affect, which were central in earlier studies of supranational identifications (e.g. Inglehart, 1970), need to be brought back to the center of explanation.
The article’s findings are also relevant for researchers who are specifically interested in the European Union and in prospects for further European integration. The statistical results above show that the impact of ML on the importance people attach to being European is not trivial. This means that the EU member states’ push to promote multilingualism since the late 1990s could contribute, if successful and followed by increased long-term mobility across the European space, to the development of a widely shared European “we-feeling” among the population. English, in particular, could play a similar role in the development of a pan-national identification to the one it has played, for instance, for India’s numerous and highly diverse ethnic groups (Laitin, 1988) and for the United States’ nationally diverse Asian community (Espiritu, 1992).

Endnotes

1. \( Q_i = p_i \times c_i = \left( \frac{P_i}{N^p} \right) \times \left( \frac{C_i}{M^p} \right) \). The prevalence of a language \( (P_i) \) is the ratio of the number of speakers of a given language by the total population in the European Union. The centrality of a language \( (C_i) \) is the ratio of the number of multilingual speakers of foreign language \( i \) by the total number of speakers of foreign languages in the European Union.

References


Boomgaarden, Hajo and André Freire. 2009. “Religion and Euroscepticism: Direct, Indirect or No Effects?” *West European Politics* 32, 6: 1240-1265


Inglehart, Ronald. 1970. “Cognitive Mobilization and European Identity.” *Comparative Politics* 3, 1: 45-70


and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, New York: Cambridge University Press


### Table 1. Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td># Languages</td>
<td>1.86</td>
<td>0.98</td>
</tr>
<tr>
<td>Speaks English (non-English speaking countries)</td>
<td>0.35</td>
<td>0.48</td>
</tr>
<tr>
<td>Importance of Being European (0-3)</td>
<td>1.64</td>
<td>0.94</td>
</tr>
<tr>
<td>Education (1-23+)</td>
<td>18.2</td>
<td>3.4</td>
</tr>
<tr>
<td>Gender (Male = 1)</td>
<td>46.7%</td>
<td>49.9%</td>
</tr>
<tr>
<td>Age (15 +)</td>
<td>47.9</td>
<td>18.1</td>
</tr>
<tr>
<td>Relative Abroad</td>
<td>29.5</td>
<td>45.6</td>
</tr>
<tr>
<td>Worked Abroad</td>
<td>11.7</td>
<td>32.1</td>
</tr>
<tr>
<td>Lived Abroad</td>
<td>7.1</td>
<td>25.6</td>
</tr>
<tr>
<td>Holidays Abroad</td>
<td>19.6</td>
<td>39.7</td>
</tr>
<tr>
<td>GDP/Capita ($)</td>
<td>31706</td>
<td>18589</td>
</tr>
<tr>
<td>Q Index</td>
<td>0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>% Protestant Population</td>
<td>22.9</td>
<td>29.7</td>
</tr>
<tr>
<td>Average Growth Rate (5-years)</td>
<td>1.71</td>
<td>1.45</td>
</tr>
</tbody>
</table>

N= 22986

Source: Eurobarometer 73.3
Table 2. The effect of the Number of Spoken Languages on European Identification (Importance attached to being European)

<table>
<thead>
<tr>
<th>Importance Attached to being European</th>
<th>OLS-Fixed Effects</th>
<th>IV-Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>SE</td>
</tr>
<tr>
<td># Languages</td>
<td>0.06</td>
<td>9.6E-03*</td>
</tr>
<tr>
<td>Education</td>
<td>0.04</td>
<td>4.4E-03*</td>
</tr>
<tr>
<td>Gender</td>
<td>-3.1E-03</td>
<td>0.02</td>
</tr>
<tr>
<td>Age</td>
<td>-0.01</td>
<td>3.3E-03*</td>
</tr>
<tr>
<td>Age Square</td>
<td>1.4E-04</td>
<td>4.0E-05*</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.85</td>
<td>0.85*</td>
</tr>
<tr>
<td>R-Squared (%)</td>
<td></td>
<td>11.1</td>
</tr>
</tbody>
</table>

2-Stage Least-Square Instrumental Variable Model (Diagnostic):
- \( r \) (Relative Abroad--Languages): 0.11 *
- Partial \( r \) (Relative Abroad -- #Languages): 0.01 *
- Partial \( b \) (Relative Abroad \( \rightarrow \) #Languages): 0.19 0.02 *

F-Value for first stage model explaining ML (Robust Standard Errors): 97.3 *
Test of Endogeneity (F-value for robust regression): 6.39 *

Source: Eurobarometer 73.3

Note: OLS and IV-Fixed Effects Models include country dummy variables. Robust Standard Errors in All Models.

\( r \) (#Languages—Importance of being European) = 0.06

*= Sig at 0.05, two-tailed
Table 3. Interaction and the Effect of Language on European Identification: The Roles of Experience Abroad, the National Language’s Communicative Value, and English

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1) OLS-Fixed Effects</th>
<th>(2) Random Slopes Model</th>
<th>(3) OLS-Fixed Effects</th>
<th>(4) OLS-Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluent in English and Other languages</td>
<td>0.09</td>
<td>0.02</td>
<td>*</td>
<td>0.03</td>
</tr>
<tr>
<td>Fluent in languages other than English</td>
<td>0.10</td>
<td>0.03</td>
<td>*</td>
<td>0.05</td>
</tr>
<tr>
<td>#Languages</td>
<td>0.04</td>
<td>9.6E-03</td>
<td>*</td>
<td>0.07</td>
</tr>
<tr>
<td>Q-index</td>
<td>-1.15</td>
<td>0.43</td>
<td>*</td>
<td>0.05</td>
</tr>
<tr>
<td>Education</td>
<td>0.04</td>
<td>4.2E-03</td>
<td>*</td>
<td>0.05</td>
</tr>
<tr>
<td>Gender</td>
<td>-6.8E-03</td>
<td>0.02</td>
<td>*</td>
<td>-5.1E-03</td>
</tr>
<tr>
<td>Age</td>
<td>-0.01</td>
<td>3.4E-03</td>
<td>-9.7E-03</td>
<td>3.2E-03</td>
</tr>
<tr>
<td>Age Square</td>
<td>1.4E-04</td>
<td>4.0E-05</td>
<td>1.2E-04</td>
<td>3.9E-05</td>
</tr>
<tr>
<td>Worked Abroad</td>
<td>0.05</td>
<td>0.03</td>
<td>*</td>
<td>0.03</td>
</tr>
<tr>
<td>Lived Abroad</td>
<td>0.03</td>
<td>0.02</td>
<td>*</td>
<td>0.03</td>
</tr>
<tr>
<td>Spends Holidays Abroad</td>
<td>0.09</td>
<td>0.02</td>
<td>*</td>
<td>0.04</td>
</tr>
<tr>
<td>Ln GDP/Cap</td>
<td>0.04</td>
<td>0.08</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Average Growth</td>
<td>0.04</td>
<td>0.08</td>
<td></td>
<td>-0.03</td>
</tr>
<tr>
<td>Square</td>
<td>0.04</td>
<td>0.08</td>
<td></td>
<td>-0.03</td>
</tr>
<tr>
<td>Root(%Protestant)</td>
<td>0.87</td>
<td>0.11</td>
<td>*</td>
<td>0.52</td>
</tr>
<tr>
<td>Constant</td>
<td>11.2</td>
<td>9.24</td>
<td>9.31</td>
<td></td>
</tr>
<tr>
<td>Random Components</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (Intercepts)</td>
<td>0.34</td>
<td></td>
<td>*</td>
<td>0.34</td>
</tr>
<tr>
<td>Variance (Slope #languages)</td>
<td>2.4E-03</td>
<td></td>
<td>*</td>
<td>2.4E-03</td>
</tr>
<tr>
<td>Variance (Slope Education)</td>
<td>6.2E-04</td>
<td></td>
<td>*</td>
<td>6.2E-04</td>
</tr>
<tr>
<td>Cov (Intercepts, Slope #languages)</td>
<td>-0.3</td>
<td></td>
<td>*</td>
<td>-0.3</td>
</tr>
<tr>
<td>Cov (Intercepts, Slope Education)</td>
<td>-0.01</td>
<td></td>
<td>*</td>
<td>-0.01</td>
</tr>
<tr>
<td>Cov (Slope #language, Slope Education)</td>
<td>9.3E-03</td>
<td></td>
<td>*</td>
<td>9.3E-03</td>
</tr>
<tr>
<td>Variance (Error)</td>
<td>0.78</td>
<td></td>
<td>*</td>
<td>0.78</td>
</tr>
<tr>
<td>N</td>
<td>22986</td>
<td></td>
<td>20620</td>
<td></td>
</tr>
</tbody>
</table>

Source: *Eurobarometer 73.3*

Note: Models 3 and 4 exclude countries where English is the official language.

Robust Standard Errors in All Models.

OLS-Fixed Effects models control for country of residence.

* = sig at 0.05, two-tailed
Table 4. The Effect of the Number of Spoken Languages, of a Language’s Communicative Value, and of Speaking English vs. Speaking other Foreign Languages on Work Abroad, Residence Abroad, and Holidays Abroad

<table>
<thead>
<tr>
<th>Variables</th>
<th>Work Abroad (1) Random Slopes Model</th>
<th>Live Abroad (2) Random Slopes Model</th>
<th>Holidays Abroad (3) Random Slopes Model</th>
<th>Work Abroad (4) OLS-Fixed Effects</th>
<th>Live Abroad (5) OLS-Fixed Effects</th>
<th>Holidays Abroad (6) OLS-Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluent in English and other languages</td>
<td>b = -7.5E-03, SE = 0.01</td>
<td>b = -5.3E-03, SE = 9.0E-03</td>
<td>b = 0.03, SE = 0.02</td>
<td>b = 3.1E-03, SE = 0.01</td>
<td>b = 0.2, SE = 0.02</td>
<td></td>
</tr>
<tr>
<td>Fluent in languages other than English #Languages</td>
<td>0.09, 6.1E-03, *</td>
<td>0.06, 4.9E-03, *</td>
<td>0.07, 5.6E-03, *</td>
<td>0.08, 0.01, *</td>
<td>0.06, 8.2E-03, *</td>
<td>0.07, 7.3E-03, *</td>
</tr>
<tr>
<td>Q-index</td>
<td>0.28, 0.07, *</td>
<td>0.18, 0.04, *</td>
<td>0.06, 0.17, *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>5.2E-03, 1.9E-03, *</td>
<td>3.1E-03, 1.1E-03, *</td>
<td>9.3E-03, 2.0E-03, *</td>
<td>3.2E-03, 1.8E-03, *</td>
<td>2.9E-03, 9.6E-04, *</td>
<td>5.5E-03, 1.5E-03, *</td>
</tr>
<tr>
<td>Gender</td>
<td>0.05, 6.3E-03, *</td>
<td>6.4E-03, 6.3E-03, *</td>
<td>7.8E-03, 8.1E-03, *</td>
<td>0.05, 6.6E-03, *</td>
<td>5.6E-03, 7.2E-03, *</td>
<td>6.7E-03, 0.08E-03</td>
</tr>
<tr>
<td>Age</td>
<td>5.4E-03, 7.8E-04, *</td>
<td>2.3E-03, 4.9E-04, *</td>
<td>4.2E-03, 9.8E-04, *</td>
<td>5.4, 9.6E-04, *</td>
<td>2.3E-03, 5.4E-04, *</td>
<td>4.2E-03, 1.0E-03, *</td>
</tr>
<tr>
<td>Age Square</td>
<td>-4.2E-05, 8.0E-06, *</td>
<td>-1.4E-05, 4.3E-06, *</td>
<td>-4.7E-05, 9.7E-06, *</td>
<td>-4.5E-05, 9.4E-06, *</td>
<td>-1.5E-05, 4.8E-06, *</td>
<td>-4.5E-05, 1.0E-05, *</td>
</tr>
<tr>
<td>Worked Abroad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lived Abroad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spends Holidays Abroad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln GDP/Cap</td>
<td>9.4E-03, 8.4E-03, *</td>
<td>1.8E-03, 6.8E-03, *</td>
<td>0.07, 0.03, *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Growth</td>
<td>1.3E-03, 4.1E-03, *</td>
<td>-3.4E-03, 2.2E-03, *</td>
<td>7.8E-03, 0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square Root(%Protestant)</td>
<td>-2.6E-03, 1.9E-03, *</td>
<td>-3.0E-03, 1.0E-03, *</td>
<td>6.6E-03, 3.7E-03, *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.40, 0.09, *</td>
<td>-0.17, 0.07, *</td>
<td>-0.94, 0.36, *</td>
<td>-0.24, 0.03, *</td>
<td>-0.12, 0.03, *</td>
<td>0.19, 0.04, *</td>
</tr>
<tr>
<td>R-Squared (%)</td>
<td>7.72, 5.37, 2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Components</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (Intercepts)</td>
<td>0.02, 7.3E-03, *</td>
<td>0.01, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (Slope #languages)</td>
<td>7.6E-04, 4.5E-04, *</td>
<td>5.4E-04, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (Slope Education)</td>
<td>6.9E-05, 2.1E-05, *</td>
<td>1.1E-04, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cov (Intercepts, Slope #language)</td>
<td>5.8E-04, -8.7E-04, *</td>
<td>-8.3E-04, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cov (Slope #language, Slope Education)</td>
<td>-1.1E-03, -3.8E-04, *</td>
<td>-1.2E-03, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cov (Slope #language, Slope Education)</td>
<td>-5.7E-05, 2.3E-05, *</td>
<td>1.2E-04, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (Error)</td>
<td>0.09, 0.06, 0.12, *</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>22986, 20620</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Eurobarometer 73.3

Note: OLS-Fixed Effects models control for country of residence; Robust Standard Errors in All Models. * = sig at 0.05, two-tailed