

What should everyone know about language? On the fluidity of important questions in linguistics

Tomas Lehecka
Åbo Akademi University

Jan-Ola Östman
University of Helsinki

Abstract

This contribution examines whether there is agreement within the global community of linguists on what should constitute common knowledge about language among the general public. We report the results of a large-scale survey study where we asked established linguists around the world ($n = 552$) to rate 15 language-related questions with respect to how important it is that the public knows the answer to them. We analyze the ratings in relation to the demographic data that we collected from the respondents. Using ordinal logistic regression models, we show that the opinions regarding what is important for everyone to know vary between linguists from different parts of the world as well as between linguists working in different subfields of linguistics. The study provides an empirical starting point for a broader reflection on the field of linguistics and the variation therein with respect to views about science communication and public outreach.

Keywords: common knowledge, discipline of linguistics, global survey, public outreach, science communication

1 Introduction

What do we linguists have to offer to the world? In addition to furthering our knowledge about language through meticulous analyses of the structure, meaning and function of various kinds of communicative behavior, we also need to inform the general public about our findings and insights. First, because the general public has, either directly (e.g. through donations) or indirectly (e.g. through taxes), supported our research; and second – and more importantly – because a deeper understanding of how language works

is crucial in close to all spheres of life. But do we know in what respects and in what areas people might need education about language? And can we linguists all agree on what would be important for everyone to know about language in the first place? Linguistics is a broad field encompassing a number of “competing” theoretical factions and methodological practices. This study therefore asks whether linguists in general – despite their different dispositions, theoretical stances and methods *with respect to research* – still share a common outlook on what constitutes the core of linguistic knowledge *for the purposes of science communication and public outreach*.

We are at present directing a large-scale research project that aims to find out what the general public knows – and does not know – about language; in other words, what linguistic realities people are (un)aware of, what misconceptions there exist, and how prevalent these misconceptions are among the general public. The project is inspired by recent efforts in several other scientific fields (e.g. medicine, sociology, environmental science) examining laypersons’ perceptions of the state of the world and their awareness of scientific facts and phenomena (e.g. Lorenzoni & Pidgeon 2006; Duffy 2018; Rosling et al. 2018; De Figueiredo et al. 2020). The ultimate aim of our project is to examine existing policies of common practices in teaching about languages and linguistics, and to suggest modifications to these practices based on our findings. However, the first step in our endeavor is to determine what the general public actually *should* know about language. To do this, we need to examine what type of knowledge and which set of facts could have the largest impact on the well-being of individuals and societies. While there may not exist any definitive, objective measure to judge the impact of individual facts, employing the judgment of a *multitude* of experts – i.e. basing the judgment on the consensus among the scientific community – seems the most appropriate way to get a handle on this issue. To this end, we carried out a global survey study among experts in linguistics asking what they perceive as the most important facts about language that everyone should be aware of.

In this contribution, we present the first results of this global survey study with 552 participating linguists from around the world. In particular, we present the results from a rating task included in the survey where the participants rated the importance of 15 language-related questions from the perspective of what should constitute common (linguistic) knowledge among the general public. The set of questions include, for example, “How many languages are there currently in the world?”, “What brain regions control

our language?”, and “How can social status and power be expressed through language?”. The main objective of the present study is to examine the ratings in relation to the demographic data that we collected from the respondents. Hence, this study seeks to answer the following primary research questions:

- i. Is there universal agreement within the linguistics community about what is important for the general public to know about language?
- ii. If not, what are some of the main demographic characteristics that predict the variation with respect to how important certain topics are perceived?

We use ordinal logistic regression models to analyze how the perceived public importance of the 15 questions varies between linguists from different parts of the world as well as between linguists working in different subfields of linguistics, while also controlling for age and gender.

To our knowledge, this is the first large-scale study of linguists’ views on what the general public should know about language. Given that the 15 questions cover only a small portion of the potentially infinite set of questions that could be asked about language, the ratings examined in this study constitute by no means an exhaustive examination of the extent to which linguists’ views on the issue vary. Rather, the study should be seen as a first exploration of the current views on the issue within the linguistics community. At the same time, the aim of this study is to inspire, and to provide an empirical foundation for, a more systematic scholarly debate on what would be good for everyone to know about language. We consider this an important debate to have for anyone interested in the value of linguistics in society in general, and for those invested in educating the public about language in particular.

2 Related work and theoretical background

Several decades ago, Richard Hudson compiled a list of statements about linguistics and language “which are likely to be accepted by virtually all linguists, irrespective of what they think about all the many issues on which linguists disagree” (Hudson 1981: 333). He compiled the list at the suggestion of the Committee for Linguistics in Education in Great Britain in order to facilitate a discussion on the relevance of linguistics to schools. The list contains 83 statements that Hudson groups thematically into three broad categories: (i) the linguistic approach to the study of language, (ii) language,

society and the individual, and (iii) language structure. The list was compiled based on several rounds of comments from altogether 46 linguists from 15 different linguistics departments in British universities. While Hudson's list shows that there are some issues in linguistics on which many linguists agree from a factual standpoint, it is unclear whether linguists also agree on how important or relevant the respective issue is for the purposes of the general public. This is a question that the present study sets out to examine.

While ours is the first survey study that deals explicitly with what linguists think everyone should know about language, one can infer some existing views on the matter from books that have been written specifically with the general public in mind (e.g. Bauer & Trudgill 1998; Crystal 2010; Napoli & Lee-Schoenfeld 2010; McWhorter 2011) as well as from linguistics-themed syllabuses designed for primary and secondary education (see e.g. Johansson et al. 2013; Lidz & Kronrod 2014; Honda & O'Neil 2017; Larson et al. 2019). However, most such resources are the product of individual authors or of relatively small working groups, leaving open the question to what extent views vary among the linguistics community at large. Furthermore, even the issues that are included in such resources may not all be considered equally important by the authors themselves. Therefore, a more systematic investigation of the current views among linguists is warranted.

Identifying what the general public should ideally know about a scientific topic makes up a crucial part of the process of science communication. Indeed, from a theoretical standpoint, we position the present study in the growing field of science communication research. According to the *mental models approach* to science communication (Bruine de Bruin & Bostrom 2013), the process of developing communication materials and activities consists of the following four components: (i) identifying what people need to know to make more informed decisions, (ii) identifying what people already know and how they make their decisions, (iii) designing communication content, and (iv) testing the effectiveness of the created content (see also e.g. Fischhoff 2013). The present study focuses on the challenges associated with the first component of this process by comparing how linguists around the world rate a number of specific questions about language in terms of how important it is that laypeople know the answer to them.

Theorists have proposed several conceptual models that represent how science is, or should be, communicated (for overviews, see e.g. Burns et al. 2003; Brossard & Lewenstein 2009). Historically, science communication has been primarily predicated on the so-called *knowledge deficit model*. Here,

science communication is driven by what experts perceive as a deficit of scientific knowledge among the target audience, and the communication is implemented as a top-down transmission of information from experts to the public. The knowledge deficit model has been widely criticized (see e.g. Sturgis & Allum 2004; Nisbet & Mooney 2007; Cortassa 2016; Simis et al. 2016) because it seems to overlook the effects of the cultural context as well as of the social and personal schemas that shape the way people interpret information. Furthermore, according to the critics of the model, science communication needs to move away from the linear top-down transmission of information and, rather, include the public in the communication as active partners in a dialogue about the perceptions, concerns and needs with regard to science (e.g. Wynne 2006; Bucchi 2008; Trench 2008). Nowadays, many science communicators also encourage scientists to include laypeople in producing new knowledge (a practice known as *citizen science*), e.g. by having them collect observational data (e.g. Bucchi & Neresini 2008; Bubela et al. 2009; Canfield et al. 2020).

It should be noted that some scholars perceive these recent trends in science communication practices as complementing, rather than replacing, the practices traditionally associated with the knowledge deficit model (e.g. Wright & Nerlich 2006; Trench 2008; Suldoovsky 2016). Testing the public's knowledge by means of large-scale surveys – a practice that has often been used in projects operating within the knowledge deficit framework – is still an essential means for determining the prevalence of different views and beliefs on scientific issues among the public (Bruine de Bruin & Bostrom 2013; Fischhoff 2013). Such surveys can reveal what themes and topics science communication will do well to focus on, irrespective of what communication strategies one considers most appropriate. Moreover, such surveys are still one of the few methods to meaningfully evaluate the outcomes of science communication efforts, i.e. to evaluate whether the target audience displays an improved understanding of a topic after having taken part in a given communication event (Bauer et al. 2007; Bauer 2009; Weingart & Joubert 2019). Naturally, the findings of survey studies always depend on what items or questions the surveys include. We see the present study as an opening in a scholarly debate on whether there exists some set of facts about language that all linguists could agree is crucial for the general public to be aware of, and that – consequently – could constitute meaningful survey items for evaluating the level of the public's knowledge about language.

3 Methods

3.1 The survey and the respondents

We collected the data for the present study using an online survey that was open for three months in the fall of 2020. The survey was hosted on the SoSci Survey platform (Leiner 2019). The survey consisted of both a rating task and open-ended questions concerning the respondents' views on what everyone should know about language. In the present study, we focus exclusively on the results from the rating task.¹ Here, we asked respondents to rate 15 language-related questions in terms of how important it is that the general public knows the answer to them. By asking all respondents to rate a fixed set of pre-formulated questions, we were able to carry out a rigorous quantitative analysis of the potential variation between the responses of different demographic groups. Figure 1 shows a screenshot of the rating task with the 15 questions. All 15 questions were presented on the same survey page, and they were presented in a random order across respondents in order to reduce potential question order bias. As the instruction sentence at the top of the figure indicates, the rating task was given after the respondents had been asked to supply their own suggestions for questions about language that everyone should know the answer to; cf. Lehecka & Östman (submitted[b]).

Given that our main objective was to determine if there exists any variation in linguists' perceptions of what is important for the general public to know about language, the set of 15 questions selected for the rating task could have been more or less arbitrary. Nevertheless, we selected the 15 questions partly based on Hudson's (1981) list of 83 issues that linguists commonly agree on (see § 2), while also selecting topics that are typically covered in introductory textbooks in linguistics (e.g. Dawson & Phelan 2016; Akmajian et al. 2017). At the same time, we strived for a reasonable spread of topics covered by the questions, and we also included questions that, based on our own intuitions, vary in terms of their importance to the general public. The ratings were collected on a three-point scale ("not so important", "quite important" and "very important"). We restricted the number of questions to 15 based on the feedback to a pilot study that we carried out among our linguistically trained colleagues ($n = 15$) at our universities (Åbo Akademi University and the University of Helsinki), where many felt that rating a larger set of questions

¹ The findings regarding the open-ended responses are reported in Lehecka & Östman (submitted[a],[b]).

23% completed

In addition to your spontaneous responses, we would appreciate it if you could rate a few sample questions regarding how important you think they are.

Please rate the questions on a 3-point scale from "Not so important" to "Very important".

	Not so important	Quite important	Very important
Do words have objectively correct meanings?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What is the relationship between standard language and dialects?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is English more closely related to Hebrew or to Persian?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What brain regions control our language?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What percentage of the world's adult population can neither read nor write?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Which three languages have the most native speakers in your country of residence?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Which six languages have the most native speakers in the world?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How many different words for 'snow' are there in the Inuit languages?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What is a morpheme?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How can social status and power be expressed through language?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What percentage of the world population has at least a basic knowledge of English?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Is sign language identical in all countries of the world?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What is the difference between a consonant and a vowel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How many languages are there currently in the world?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What portion of the world's languages typically place the subject in front of the verb?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1. A screenshot of the survey page with the rating task

was too tedious. Similarly, we chose to use a three-point scale based on the feedback to the pilot study, where respondents argued that rating questions on a more fine-grained scale was too difficult and too time-consuming.

In order to gather responses from as many linguists as possible, we sent out survey invitations to the linguistics departments of the 300 highest ranked universities on the *QS World University Rankings of Linguistics (QS 2020)* asking them to distribute the survey link among their academic staff and associates working in linguistics and language studies. Also, we made the survey link available through *LinguistList* as well as through many other linguistics fora (e. g. different national and international associations). In total, we received over 600 (anonymous) responses to the survey. Out of these, we included the responses from 552 respondents for detailed analysis. Based on

the background information we collected from the respondents (in particular, their academic rank, affiliation, and what linguistic subfields they work in), we feel confident that these 552 respondents are, indeed, established academics in the field of linguistics, scholars with expert knowledge in linguistics or language studies, or they are employed in a linguistic profession. By contrast, we excluded those respondents from the sample who had indicated that they are undergraduate students or that they do not work professionally with language.

The sample of respondents is geographically diverse. All continents except Antarctica and at least 51 different countries of residence are represented in the sample (23 respondents preferred not to disclose their country of residence). The largest number of respondents comes from the U.S.A. (82), followed by Germany (56) and the U.K. (48). At the opposite end of the spectrum, countries like Algeria, Kazakhstan, Kuwait, Macao, and Uruguay are represented only by single individuals. There is a certain over-representation of respondents from the countries where we are ourselves active (viz., Sweden, Finland, and Switzerland). It should also be noted that the number of responses from South America, Asia and Africa is relatively low, and that most of the respondents from Africa come from Nigeria. Additionally, our sample is probably influenced by the fact that the language of the survey was English; we are aware that some communities may rather favor participating in studies in their own (national) language or in the language most often used for studies in their subfield of linguistics.

The sample is also diverse with respect to the respondents' affiliations. The sample includes respondents from at least 247 different universities or institutes (105 respondents preferred not to disclose their affiliation), with Stockholm University being the most well-represented university in the sample (15 respondents). With respect to academic rank, the largest respondent groups in our sample are PhD candidates (124)², assistant or associate professors (110), and full professors (101). The sample includes more female (320) than male (193) respondents, with 11 respondents indicating other gender identity, and 28 respondents preferring not to indicate their gender. Detailed tables showing the distribution of the respondents according to the demographic variables can be found in Appendix A.

When choosing what demographic variables to collect from the

² We included PhD candidates in the sample based on the fact that PhD candidates publish academic research (thus contributing e.g. to the QS-ranking of their affiliation) and that being a PhD candidate is a full-time profession in many countries.

respondents, we restricted the selection to variables with a relatively straight-forward structure (in terms of variable levels or values) that are easily applicable in the global context (gender, age, etc.). For this reason, we left out some other “typical” demographic variables (such as class or race) which are relatively context-dependent. We also did not ask how much experience the respondents have with science education or science communication, partly because we assume that many respondents would have a hard time quantifying their experiences, and partly because scholars have different views on what counts as science communication (see § 2, see also Burns et al. 2003).

Because there exists no reference data regarding the demographics of the global population of linguists, we cannot make claims as to the representativeness of our sample. However, in view of the large number of responses and their considerable dispersion in terms of affiliations, we do feel confident that our data can be used to draw some preliminary conclusions about the “general feeling in the field” and about some of the variation that the field includes. Naturally, we regard the very willingness of these 552 linguists to participate in our study as indicating that they are interested in the topic of the survey, and thus, in communicating findings about language to the general public.

3.2 Regression analysis and categorizations

In order to examine how the perceived public importance of certain questions in linguistics co-varies with specific background variables relating to the linguists-as-respondents, we analyzed the responses to our survey by means of ordinal logistic regression models. These allow the estimation of the effect of multiple individual variables on the ratings even though the distribution of the respondents across the different levels/values of those variables is uneven (see § 3.1 and Appendix A). We built the models using the *ordinal* package (Christensen 2015) in R (R Core Team 2020). For each question, we built a model where the ratings of the perceived importance of the question constituted the dependent ordinal variable, while gender, age, region and linguistic subfield affiliation constituted the independent variables. Given that this is the first study on linguists’ views of what is important for everyone to know about language – and given that our aim is to determine whether there exists variation in these views between different groups of linguists – we chose to include such independent variables in the regression models that, based on other lines of research, give us a reasonable chance to find significant effects

on the ratings. In what follows, we briefly present the motives for including the four variables in question (gender, age, region and subfield affiliation) as independent variables in the regression analysis, and we also describe how we operationalize these variables in the analysis.

Gender and *age* are two of the most well-researched variables in human science. As for the former, there is a plethora of evidence from psychology indicating that there are significant gender differences with respect to what interests people have (e.g. Su, Rounds & Armstrong 2009; Lippa 2010), and it is plausible that different interests may translate into different perceptions of what constitutes important knowledge. In the regression models, we operationalize gender as a binary categorical variable (female/male; subjects who indicated a different gender identity were not included in the models due to their low number in the data). Similarly, age has been shown to have a significant effect on people's interests, attitudes and behavior (e.g. Twenge, Campbell & Freeman 2012; Oh & Reeves 2014); therefore, it is reasonable to assume that age could also influence people's views on what should be part of common knowledge. However, in the context of the present study, age should not be regarded as merely a biological/cognitive variable, but also as a linguistics-cultural one; a linguist's age may reflect the theoretical, methodological and thematic trends in linguistics that their thinking has been influenced by during their careers (for overviews on trends in linguistics, see e.g. Robins 1997; Newmeyer 2019). Such trends may then, in turn, influence linguists' perceptions of what constitutes core linguistic knowledge. In the regression models, we treat age as a continuous variable.

In addition to linguistic trends changing and evolving with time, different linguistic trends are (and have been) dominant in different parts of the world (see e.g. Koerner & Asher 1995). We include the variable *region* in the regression analysis in order to explore potential regional differences regarding the perceived importance of the questions in the rating task. Specifically, this variable refers to what part of the world a respondent resides in. Due to the considerable geographical dispersion of the respondents (see § 3.1), we grouped the respondents' countries of residence into three categories for the purposes of the regression models. These categories are: (i) the Anglosphere (encompassing the U.S.A., U.K., Ireland, Canada, Australia and New Zealand), (ii) Europe (excluding the U.K. and Ireland), and (iii) Other regions. Our categorization is motivated partly by theoretical reasons (the Anglosphere has to a certain extent been influenced by different theoretical currents than continental Europe; see e.g. Joseph 2002), partly by the fact

that the main language in the Anglosphere is the same as the language of the survey (i.e. English), and partly by the number of respondents from each region. Unfortunately, the category Other regions is both the smallest in terms of the number of respondents and the most heterogeneous in terms of the geographical area it covers. Therefore, we do not make any claims as to the representativeness of this category but treat it merely as an interesting point of comparison for the other two, more well-represented regions. We also carried out regression analyses to check for potential differences in the ratings by linguists from the U.S.A. and the U.K. (the two most well-represented countries of residence within the Anglosphere category). We acknowledge, of course, that a person's country of residence does not always coincide with where they have received their linguistic training, but we think this variable can, nonetheless, be regarded as a broad approximation of a respondent's linguistics-cultural background.

Finally, the variable *subfield affiliation* refers to what subfield(s) of linguistics a respondent is actively working in. Different subfields of linguistics have, almost per definition, different priorities with respect to linguistic methods and theory, which is why subfield affiliation seems an obvious choice to include in the current analysis. In the survey, respondents were asked to indicate their fields of research/work from a list of 29 linguistic subdisciplines used by the Open Language Archives Community (OLAC). Respondents were free to choose more than one subdiscipline (the median number of subfields selected was 3), and they could also write in additional subdisciplines if they felt that the pre-set choices were too restrictive. While the OLAC list of linguistic subfields can no doubt be criticized for some of its categorization, the list provides a useful, independent baseline to work with. Of OLAC's 29 subdisciplines, sociolinguistics (173), applied linguistics (162), and language acquisition (119) were checked off by the largest numbers of respondents in our sample. By contrast, fields such as history of linguistics (7), mathematical linguistics (7), and forensic linguistics (5) were indicated by only a few respondents each. For the full table indicating subfield affiliations, see Appendix A.

For the purposes of the regression analysis, we needed to operationalize subfield affiliation as a manageable number of macro-fields represented by a comparable number of respondents. However, we did not just want to rely on our own intuitions regarding how the different subfields relate to each other. Therefore, we asked for help from five internationally well-known professors of linguistics (from five different countries) who were kind enough

to suggest groupings of the 29 subfields into four to six macro-fields which they considered appropriate. Comparing the different groupings, we could conclude that, for the most part, the five individuals were relatively unanimous in what fields bear the most resemblance to each other. We then used these groupings as a starting point to create five macro-field categories which we used in the regression analysis. In order to further validate this grouping, we carried out a hierarchical cluster analysis of the subfields based on the survey responses; specifically, we clustered the subfields based on the likelihood that respondents represent subfield A, given that they also represent subfield B (more accurately, we used the Jaccard similarities between the binary yes/no-responses concerning whether respondents work in given fields). The resulting clusters provided additional support for the soundness of the macro-field categorization we used in the regression models (especially in instances where the groupings made by the five professors differed from one another), while also minimizing the risk of potential high associations between the individual subfield variables causing problems for the regression models.

The labels we give the five macro-fields in our analysis are: General (encompassing, among others, general linguistics, typology, and language documentation), Grammar (phonology, morphology, syntax, etc.), Function (sociolinguistics, pragmatics, discourse analysis, and anthropological linguistics), Cognition (psycholinguistics, neurolinguistics, language acquisition, etc.), and Applied (applied linguistics, translating and interpreting, lexicography, etc.); see Appendix B for the grouping of all subdisciplines. Obviously, these macro-fields are not mutually exclusive categories; i.e. a person can be active in several macro-fields. Therefore, in the regression models, we operationalize subfield affiliation as a set of five binary variables (using the values “yes”/“no” for each of the five macro-fields), and we apply Bonferroni correction to the p -values concerning the effects of these macro-fields on the ratings of the questions.

Due to space constraints, we report the results of the regression models in a single table (Table 1), and we only report statistically significant effects of the independent variables on the ratings. The results reflect values obtained from regression models that do not include any interaction terms between the independent variables. We think including interactions in the models is unwarranted for two reasons. First, there exists no prior evidence pointing to potential significant interactions of the independent variables with respect to the perceived importance of a question or a topic. Second, and more importantly, the variables region and subfield affiliation are constructs created

Table 1. Summary of the results of the ordinal logistic regression models

#	Question	Mean (n)	Age	Gender	Region	Linguistic macro-fields				
						General	Grammar	Function	Cognition	Applied
1	How can social status and power be expressed through language?	2.67 (552)		F > M (.007)				Y > N (.012)		
2	What is the relationship between standard language and dialects?	2.64 (550)			A > O (.032)		Y > N ($<.001$)		Y < N (.002)	
3	Do words have objectively correct meanings?	2.52 (547)								
4	Is sign language identical in all countries of the world?	2.38 (549)			A, O > E (.014)					
5	Which three languages have the most native speakers in your country of residence?	2.34 (551)			A, O < E (.002)					Y > N (.014)
6	What percentage of the world's adult population can neither read nor write?	2.32 (552)			A < E (.027) [†]					
7	What is the difference between a consonant and a vowel?	2.16 (550)			A < E (.003)					
8	How many languages are there currently in the world?	2.12 (550)		F > M (.004)						
9	Which six languages have the most native speakers in the world?	2.02 (550)			A < E (.035)					
10	What brain regions control our language?	1.80 (551)			A, E < O ($<.001$)					
11	What percentage of the world population has at least a basic knowledge of English?	1.77 (549)			A < O (.004)	Y < N (.008)		Y > N (.015)		
12	What is a morpheme?	1.72 (550)								
13	Is English more closely related to Hebrew or to Persian?	1.50 (549)								Y < N (.010)
14	What portion of the world's languages typically place the subject in front of the verb?	1.41 (550)			A, E < O ($<.001$)					
15	How many different words for 'snow' are there in the Inuit languages?	1.39 (548)		F > M (.020)						

Questions are ordered according to their average perceived importance. "Mean" represents the average rating when the ordinal variable is transposed to an integer scale of 1–3. Region codes: "A" = "the Anglosphere", "E" = "Europe", "O" = "Other regions". For linguistic macro-fields, "Y" = "Yes, subject indicates they represent at least one subfield in the macro-field", "N" = "No, subject does not represent any subfield in the macro-field". Values in brackets represent p-values obtained by means of Wald tests, except for the variable region, for which p-values were obtained by a likelihood ratio test of the full model with the effect of region against a model without that effect. P-values concerning the linguistic macro-fields are Bonferroni-corrected for multiple comparisons.

[†]Within the Anglosphere, respondents from the U.S.A. give significantly lower ratings than respondents from the U.K. ($p = .012$).

for the purposes of the present analysis and they have a very complex internal structure (see above). Interpreting potential interaction effects between such variables would be cumbersome and of questionable value in view of our research questions (see § 1). When we describe the effects of the independent variables on the perceived importance of a question (in § 4), we often refer to the predicted probability of a given outcome (e.g. the predicted probability of question (1) being rated as “very important” by female linguists). The predictions are generated by a regression model by holding the non-focal independent variables (in this case, age, region, and subfield affiliation) constant while varying the focal independent variable (in this case, gender). We used the *ggemmeans* function from the R package *ggeffects* (Lüdtke 2018) to estimate the predicted probabilities.

4 Results

Table 1 gives a summary of the results regarding the ratings of the 15 questions in our survey. In the table, the questions are ordered according to how important, on average, they were considered to be by our respondents. Before examining how the views about the importance of the individual questions vary between different respondent groups, it is worth making a few short observations about the ratings of the questions in general.

Of the 15 questions, the questions that were considered most important by the respondents are (1) “How can social status and power be expressed through language?” (73% of all participants rated this question as “very important”), (2) “What is the relationship between standard language and dialects?” (68%), and (3) “Do words have objectively correct meanings?” (62%). It may be noted that these questions have also for decades occupied a prominent position within linguistic research, which speaks to their complex nature. For example, the details of the relationship between social status, power, and language (as referred to in question (1)) are still widely debated. Similarly, how linguists answer question (2) may vary quite substantially depending on what theoretical framework they happen to be working within. As for question (3), linguists may be fairly unanimous in giving a negative answer to this question, but that answer inevitably requires a fair amount of non-trivial explication and elaboration. It seems, then, that topics that linguists consider to be of the highest public importance are also topics that are at the forefront of much of ongoing research in linguistics.

By comparison, the three questions that were rated as least important by our respondents are (13) “Is English more closely related to Hebrew or to Persian?” (only 9% of the participants rated this question as “very important”), (14) “What portion of the world’s languages typically place the subject in front of the verb?” (5%), and (15) “How many different words for ‘snow’ are there in the Inuit languages?” (5%). The fact that question (15) received low ratings is understandable from a theoretical point of view – after all, the lexicon in Inuit languages has very little impact on most people’s lives – but at the same time, this is the kind of question that, based on anecdotal evidence, typically interests the general public. In a similar vein, the fact that the respondents felt that question (14), despite its unquestionable position in all traditional textbooks on linguistics, is not very important for the general public to know about may be taken as an indication that the respondents were, indeed, thinking outside of the purely linguistic perspective when rating the questions. As for question (13), evaluating the importance of knowing about the historical relationships between languages is probably influenced by which languages one specifically mentions. In some ways, however, this question is similar to question (15) in the sense that etymology and the historical relationships between languages tend to be quite salient in laypeople’s minds when they talk to linguists (Stollznow 2018), but they are not prioritized by linguists for the purpose of the issues in this study.

After these general observations, we now move on to examining how the ratings of the 15 questions co-vary with the background variables relating to the respondents. Table 1 illustrates, separately for each question, the results of the ordinal logistic regression models. We only report statistically significant effects in the table. For these effects, the table indicates what variable levels predict lower or higher ratings of the question when all the other variables are kept constant. As Table 1 demonstrates, we find significant effects of the independent variables on the ratings for all but two of the questions. With respect to questions (3) and (12), there are no statistically significant differences in the ratings that are due to the background variables included in the models. For all other questions, the ratings co-vary with at least one of the independent variables. In what follows, we describe the effects in detail separately for each variable.

Age

We do not find any significant effects of age on the ratings. This means that none of the questions was rated as more important by younger linguists compared to older linguists, or vice versa. Given that the rating task only consisted of 15 questions, we *cannot* conclude that age *never* plays a role in linguists' perceptions of what constitutes important linguistic knowledge. For example, had the task included a question about new technologies or about a linguistic theory associated with a specific time period, it is not inconceivable that we might see an effect of age on the ratings. Nevertheless, the results of our study strongly suggest that, with respect to most issues, linguists' perceptions are not influenced by age – especially when compared to gender, region and subfield affiliation.

Gender

We find a significant effect of gender on the ratings of three of the questions: (1) “How can social status and power be expressed through language?”, (8) “How many languages are there currently in the world?”, and (15) “How many different words for ‘snow’ are there in the Inuit languages?”. Specifically, female respondents rated these questions as more important than what male respondents did. Based on the regression model for question (1), the predicted probability of a female linguist rating this question as “very important” is 75%, while the probability of a male linguist doing so is 65%. This difference may not be surprising from an historical or feminist perspective, but we think it is important to recognize that such a difference still exists in the implicit attitudes of male and female linguists. In short, the issue of language and power, and its implications for the general public, are prioritized by women to a greater degree than by men.

In comparison with question (1), the gender effects concerning questions (8) and (15) are perhaps less expected. The predicted probability of a female linguist rating question (8) as “very important” is 44%, while the probability of a male linguist doing so is 32%. Question (15) received relatively low ratings overall, but the predicted probability of a female linguist rating it as at least “quite important” is 40%, while the probability of a male linguist doing so is 28%. Our data do not tell us why these gender differences exist, nor why we find such a difference for question (8) but not, for example, for questions (5) and (9), given that they concern relatively similar topics. Be that as it may,

we think it is crucial to recognize that gender is a factor in relation to what linguists consider important for the general public to know about language.

Region

We find significant effects of region for nine of the 15 questions. In this sense, region is the most “significant” of our background variables, i.e. region is the variable that most often yields a significant effect on the ratings of these questions. As we noted earlier (see § 3.2), due to the uneven geographical distribution of the respondents, the most interesting and reliable effects of region have to do with differences between the Anglosphere and Europe. Therefore, we describe these effects first. After that, we turn to effects that pertain to comparisons with Other regions.

To begin with, respondents from the Anglosphere (and Other regions) rated question (4) “Is sign language identical in all countries of the world?” as more important than what respondents from Europe did. Based on the regression model, the predicted probability of a respondent from the Anglosphere rating this question as “very important” is 54%, while the probability of a respondent from Europe doing so is 42%. By contrast, questions (5), (6), (7), and (9) were rated as more important by linguists based in Europe than by those based in the Anglosphere. For question (5), “Which three languages have the most native speakers in your country of residence?”, the predicted probability of a respondent from the Anglosphere rating this question as “very important” is 39%, while the probability of a European linguist doing so is 53%. For question (6), “What percentage of the world’s adult population can neither read nor write?”, the predicted probability of a respondent from the Anglosphere rating this question as “very important” is 35%, while the probability of a European linguist doing so is 47%. For question (7), “What is the difference between a consonant and a vowel?”, the corresponding predicted probabilities are 29% and 44%, and for question (9), “Which six languages have the most native speakers in the world?”, the corresponding probabilities are 21% and 29%, respectively.

We acknowledge, of course, that the notion of “the Anglosphere” is not a linguistics-theoretically homogeneous grouping. This is why we also carried out an additional ordinal logistic regression analysis of our data aiming at identifying possible differences in the ratings between respondents from the U.S.A. and the U.K., the two largest countries (as far as the number of respondents goes) within the Anglosphere. Interestingly, here we found a

significant difference for only one of the questions, namely, question (6) “What percentage of the world’s adult population can neither read nor write?”, but this difference is substantial. The predicted probability of a respondent from the U.K. rating this question as “very important” is almost twice as high (48%) as that of a respondent from the U.S.A. doing so (27%).

With respect to questions (2), (10), (11), and (14), the significant effect of region concerns the difference between the Anglosphere (and sometimes Europe) on the one hand, and Other regions on the other hand. As we noted in § 3.2, the number of respondents from Other regions is small and very unevenly distributed, which is why the effects concerning this category should be considered tentative. They do, however, provide some additional perspective on the ratings from the Anglosphere and Europe, and hopefully these tentative results will inspire future confirmatory studies with a higher number of subjects from Africa, Asia and South America. It is, for example, intriguing that the relationship between standard language and dialects was judged to have more public importance by respondents from the Anglosphere than by respondents from Other regions (the predicted probability of a linguist from the Anglosphere rating question (2) as “very important” is 71%, whereas the probability of a linguist from Other regions doing so is 53%). Similarly, it is intriguing that the results suggest a substantial difference in the opposite direction for question (10) “What brain regions control our language?” in that respondents from Other regions rated this question as more important than what respondents from the Anglosphere or from Europe did (the predicted probabilities for rating this question as “very important” are 14% for respondents from the Anglosphere, 11% for respondents from Europe, and 30% for respondents from Other regions). An even larger difference is found for question (14) “What portion of the world’s languages typically place the subject in front of the verb?”, where the predicted probabilities of being rated as at least “quite important” are 28% for linguists from the Anglosphere, 33% for linguists from Europe, and 62% for linguists from Other regions.

Subfield affiliation

We find significant effects of subfield affiliation for five of the 15 questions. In the regression models, comparisons are made between those respondents who work in a given macro-field and those who do not (see § 3.2 for details). Looking at the effects of subfield affiliation in Table 1, it appears that some are

in line with what one might expect based on the traditional foci of the respective macro-fields. For example, it is not surprising that respondents who work in the macro-field Function rated question (1) “How can social status and power be expressed through language?” as more important than what other respondents did. After all, the macro-field Function does encompass such subdisciplines as sociolinguistics, discourse analysis and pragmatics, which are traditionally interested in issues of status and power. According to the regression model for question (1), the predicted probability of a linguist working in the macro-field Function rating this question as “very important” is 78%, while the probability of linguists not working in this field doing so is 62%.

Similarly, the effects of subfield affiliation on the ratings of question (2) “What is the relationship between standard language and dialects?” are not totally unexpected. Here, respondents who work in the macro-field Cognition rated the question as less important than what respondents not working in this field did (the predicted probability of a linguist working in Cognition rating this question as “very important” is 56%; for others 70%). This is understandable given that the subdisciplines in this macro-field (such as psycholinguistics, neurolinguistics and language acquisition) traditionally have not primarily focused on issues of social prestige or language planning. At the same time, respondents who work in the macro-field Grammar (encompassing subdisciplines like phonetics, phonology, morphology, syntax and semantics) rated question (2) as more important than what respondents not working in this field did. The predicted probability of a linguist working in Grammar rating this question as “very important” is 72%, while the probability of other linguists doing so is 55%. Again, this is not entirely unexpected in light of what role concepts like grammaticality and variation occupy within these subdisciplines. We do find it interesting, however, that there is an effect of the macro-field Grammar for question (2), rather than an effect of the macro-field Function, even though this question is implicitly concerned with language and power, i. e. a topic similar to the one addressed in question (1).

With respect to question (5) “Which three languages have the most native speakers in your country of residence?”, we find that linguists working in the macro-field Applied (encompassing e. g. applied linguistics, translating and interpreting, and lexicography) rated the question as more important than what linguists not working in this field did. The predicted probability of a respondent working in the macro-field Applied rating this question as “very important” is 50%, while the probability of other respondents doing so is 37%.

Perhaps this result reflects the focus by applied linguists on real-life problems such as communication and education in multilingual settings. However, if that is the case, then it is slightly surprising that we do not also find a similar effect with respect to e.g. question (11) “What percentage of the world population has at least a working knowledge of English?”, given that it, too, concerns multilingualism (but from a global perspective). Instead, for question (11), we find that respondents working in the macro-field Function rated the question as more important than what other respondents did (the predicted probability of a linguist working in the macro-field Function rating this question as “very important” is 21%; for others 9%). In addition, respondents working in the macro-field General rated question (11) as less important than what other respondents did (the predicted probability of a linguist working in the macro-field General rating this question as at least “quite important” is 54%, while the probability of other linguists doing so is 69%).

Last, with respect to question (13) “Is English more closely related to Hebrew or to Persian?”, we find that respondents who work in the macro-field Applied rated this question as less important than what respondents not working in this field did. The predicted probability of a respondent working in Applied rating this question as at least “quite important” is 35%, while the probability of other respondents doing so is 50%. We find this effect, much like the one regarding question (11), relatively unexpected.

5 Discussion

In this study, we have examined whether there is universal agreement within the linguistics community about what is important for the general public to know about language. The data we have presented in this study are the first of their kind in that they come from an empirical, large-scale survey to the international community of linguists. In the survey, we asked respondents to rate 15 language-related questions in terms of their public relevance. The 15 questions cover only a small portion of the topics that linguistics deals with, but even this limited sample of questions is enough to demonstrate that there is disagreement among linguists regarding what is important for the general public to know. We have analyzed the variation in linguists’ views on this issue by means of ordinal regression models, which show that the variation can – at least to a certain extent – be explained (i.e. statistically predicted) with reference to certain demographic variables. Specifically, the perceived

importance of a question often varies between linguists from different parts of the world as well as between linguists working in different subfields of linguistics. For a few questions, we also find a significant effect of gender. By contrast, we do not find any effect of age on the perceived importance of the questions included in our study.

We consider the present study a first exploration of the present-day opinions within the global linguistics community regarding what should constitute common knowledge about language. Implicitly then, the study raises questions about what (current and future) efforts in science communication and public outreach in linguistics should focus on, and whether linguistics should be striving toward a unified view on this issue. Linguistics today is a very broad field, and the “outer borders” are not easy to draw against e. g. communication, education, sociology and psychology. The 552 respondents in our survey study have all come forth as linguists, albeit that their research focus and interests vary considerably. It is clear that linguistics encapsulates an abundance of different theories and views, but, for the purposes of public outreach, we might still like to see that there is a common understanding of what is important in linguistics, what our major findings are, and what facts and realities the general public should be aware of.

Further research is needed to determine the underlying reasons for the differences we find between the views of different respondent groups regarding what is important for everyone to know about language. After all, having knowledge of these reasons is a prerequisite for being able to judge to what extent the variation in linguists’ views on this issue is something that should be remedied, or whether such variation should, in fact, be seen as something positive. If the variation is simply the result of differing theoretical or linguistic traditions that are prevailing in different geographical regions or subfields of linguistics, one might argue that a more unified view of what counts as important knowledge about language should be pursued, putting what is most beneficial for the general public at the center when planning science communication and public outreach. This conclusion would be especially easy to draw if one were to adhere to the knowledge deficit model of science communication, which assumes that there is a set of basic facts that everyone should be aware of, irrespective of specific geographical or social contexts (see § 2). If, however, the differences in linguists’ views reflect that the general public has varying needs for language-related knowledge in different parts of the world or in different sub-groups of the public, the diverse views of linguists about what the general public should know about

language can be seen as benefitting the discipline. This conclusion would be in accordance with models of science communication that stress the role of social context. We do not know how the linguists that participated in our survey think about science communication from a theoretical perspective, nor what opinions they have on the practices of science education or public outreach in general (e.g. top-down vs. dialogic communication). We also do not know who they were primarily thinking of as the target audience for the questions in the rating task: everyone in the world/in their own country/in a specific community? Concepts such as ‘the general public’ or ‘knowledge’ are complex (see e.g. Burns et al. 2003), and large-scale quantitative studies such as the present one cannot take all this complexity into consideration.

It is important to acknowledge that the results presented in this study concern only the 15 questions included in the rating task, i.e. the results cannot be generalized to other questions about language. It is also likely that the ratings of perceived importance given by the linguists (with regard to the individual questions) were influenced by how the questions were worded and framed. Similarly, it is possible that the overall findings, in terms of what demographic variables are significant predictors for the ratings of the largest number of questions, would look different had we selected a different sample of questions for the rating task. Furthermore, adding a larger number of demographic variables (or variable levels) might reveal different or additional patterns regarding linguists’ views about the perceived public importance of the questions. From the standpoint of the science communication framework, one variable worth examining by future research is the level of experience (and/or interest) that the respondents have in science communication and public outreach. We think such a variable is difficult to operationalize in a short global survey such as ours (as we argue in §3.1), but it could be worthwhile to include it in more qualitatively oriented follow-up studies with in-depth interviews of a smaller sample of linguists.

Finally, we do want to argue that the present study demonstrates that surveys investigating one’s own field of science are important as they reveal that the dominant opinions and priorities in the field might not always be what we think they are. We hope this study may inspire not only linguists, but scholars in other fields as well, to discuss what type of knowledge about language could have the largest positive impact on the well-being of individuals and societies.

Acknowledgements

The research project *Consequential misconceptions about local and global linguistic realities* is directed by Tomas Lehecka and funded by The Emil Aaltonen Foundation, The Society of Swedish Literature in Finland, The Finnish Society of Sciences and Letters, and The Swedish Cultural Foundation in Finland.

We wish to express our deep gratitude to two anonymous referees for their detailed comments and suggestions to an earlier version of this paper, and to the editorial team of the *Finnish Journal of Linguistics* for their guidance throughout the review process. We would also like to extend our sincere appreciation to all those linguists who took the time to participate in our survey.

References

- Akmajian, Adrian & Farmer, Ann K. & Bickmore, Lee & Demers, Richard A. & Harnish, Robert M. 2017. *Linguistics: An introduction to language and communication*. Cambridge, MA: MIT Press.
- Bauer, Laurie & Trudgill, Peter (eds.). 1998. *Language myths*. London: Penguin.
- Bauer, Martin W. 2009. The evolution of public understanding of science: Discourse and comparative evidence. *Science, Technology and Society* 14(2). 221–240.
- Bauer, Martin W. & Allum, Nick & Miller, Steve. 2007. What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public Understanding of Science* 16(1). 79–95.
- Brossard, Dominique & Lewenstein, Bruce V. 2009. A critical appraisal of models of public understanding of science: Using practice to inform theory. In Kahlor, LeeAnn & Stout, Patricia A. (eds.), *Communicating science*, 25–53. New York: Routledge.
- Bruine de Bruin, Wändi & Bostrom, Ann. 2013. Assessing what to address in science communication. *Proceedings of the National Academy of Sciences* 110(supplement_3). 14062–14068.
- Bubela, Tania & Nisbet, Matthew C. & Borchelt, Rick & al. [= 24 authors]. 2009. Science communication reconsidered. *Nature Biotechnology* 27(6). 514–518.
- Bucchi, Massimiano. 2008. Of deficits, deviations and dialogues: Theories of public communication of science. In Bucchi, Massimiano & Trench, Brian (eds.), *Handbook of public communication of science and technology*, 71–90. London: Routledge.

- Bucchi, Massimiano & Neresini, Federico. 2008. Science and public participation. In Hackett, Edward J. & Amsterdamska, Olga & Lynch, Michael & Wajcman, Judy (eds.), *The handbook of science and technology studies*, 449–472. 3rd edn. Cambridge, MA: MIT Press.
- Burns, Terry W. & O'Connor, D. John & Stocklmayer, Susan M. 2003. Science communication: A contemporary definition. *Public Understanding of Science* 12(2). 183–202.
- Canfield, Katherine N. & Menezes, Sunshine & Matsuda, Shayle B. & al. [= 12 authors]. 2020. Science communication demands a critical approach that centers inclusion, equity, and intersectionality. *Frontiers in Communication* 5(2).
- Christensen, Rune Haubo B. 2015. *Ordinal: Regression models for ordinal data*. (R package version 2015.6-28). (<http://www.cran.r-project.org/package=ordinal/>).
- Cortassa, Carina. 2016. In science communication, why does the idea of a public deficit always return? The eternal recurrence of the public deficit. *Public Understanding of Science* 25(4). 447–459.
- Crystal, David. 2010. *The Cambridge encyclopedia of language*. 3rd edn. Cambridge University Press.
- Dawson, Hope & Phelan, Michael (eds.). 2016. *Language files: Materials for an introduction to language and linguistics*. Columbus: The Ohio State University Press.
- De Figueiredo, Alexandre & Simas, Clarissa & Karafillakis, Emilie & Paterson, Pauline & Larson, Heidi J. 2020. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: A large-scale retrospective temporal modelling study. *The Lancet* 396(10255). 898–908.
- Duffy, Bobby. 2018. *The perils of perception: Why we're wrong about nearly everything*. London: Atlantic Books.
- Fischhoff, Baruch. 2013. The sciences of science communication. *Proceedings of the National Academy of Sciences* 110(supplement_3). 14033–14039.
- Honda, Maya & O'Neil, Wayne. 2017. On thinking linguistically. *Revista Linguística* 13(1). 52–65.
- Hudson, Richard. 1981. Some issues on which linguists can agree. *Journal of Linguistics* 17(2). 333–343.
- Johansson, Victoria & Carling, Gerd & Holmer, Arthur & al. [= 30 authors]. 2013. *Språket, människan och världen*. Lund: Studentlitteratur.
- Joseph, John E. 2002. *From Whitney to Chomsky: Essays in the history of American linguistics*. Amsterdam: John Benjamins.
- Koerner, Ernst F. K. & Asher, Roland E. (eds.). 1995. *Concise history of the language sciences*. New York: Elsevier.
- Larson, Richard & Denham, Kristin & Lobeck, Anne. 2019. The AP Linguistics initiative. *Language* 95(3). e381–e393.

- Lehecka, Tomas & Östman, Jan-Ola. Submitted(a). Toward establishing what linguists think the general public should know about language: Comparing open-ended responses and rating questions.
- Submitted(b). What linguists think everyone should know about language: Analyzing responses to open-ended questions.
- Leiner, Dominik J. 2019. *SoSci Survey (version 3.1.06)*. (<https://www.soscisurvey.de/>).
- Lidz, Jeffrey & Kronrod, Yakov. 2014. Expanding our reach and theirs: When linguists go to high school. *Language and Linguistics Compass* 8(10). 449–463.
- Lorenzoni, Irene & Pidgeon, Nick F. 2006. Public views on climate change: European and USA perspectives. *Climatic Change* 77(1). 73–95.
- Lüdecke, Daniel. 2018. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* 3(26). 772. DOI: 10.21105/joss.00772.
- McWhorter, John. 2011. *What language is (and what it isn't and what it could be)*. New York: Gotham Books.
- Napoli, Donna Jo & Lee-Schoenfeld, Vera. 2010. *Language matters: A guide to everyday questions about language*. 2nd edn. New York: Oxford University Press.
- Newmeyer, Frederick J. 2019. Martin Joos's readings in linguistics: A publication history. *Historiographia Linguistica* 46(3). 73–130.
- Nisbet, Matthew C. & Mooney, Chris. 2007. Framing science. *Science* 316(5821). 56.
- QS. 2020. = *QS world university rankings for linguistics*. (<https://www.topuniversities.com/university-rankings/university-subject-rankings/2020/linguistics>).
- R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (<https://www.R-project.org/>).
- Robins, Robert Henry. 1997. *A short history of linguistics*. 4th edn. London: Routledge.
- Rosling, Hans & Rosling, Ola & Rosling Rönnlund, Anna. 2018. *Factfulness: Ten reasons we're wrong about the world – and why things are better than you think*. New York: Flatiron Books.
- Simis, Molly J. & Madden, Haley & Cacciatore, Michael A. & Yeo, Sara K. 2016. The lure of rationality: Why does the deficit model persist in science communication? *Public Understanding of Science* 25(4). 400–414.
- Stollznow, Karen. 2018. “How many languages do you speak?” Perceptions and misconceptions about linguistics and linguists. *Lingua* 205. 15–28.
- Sturgis, Patrick & Allum, Nick. 2004. Science in society: Re-evaluating the deficit model of public attitudes. *Public Understanding of Science* 13(1). 55–74.
- Suldovsky, Brianne. 2016. In science communication, why does the idea of the public deficit always return? Exploring key influences. *Public Understanding of Science* 25(4). 415–426.

- Trench, Brian. 2008. Towards an analytical framework of science communication models. In Cheng, Donghong & Claessens, Michel & Gascoigne, Toss & Metcalfe, Jenni & Schiele, Bernard & Shi, Shunke (eds.), *Communicating science in social contexts*, 119–135. New York: Springer.
- Weingart, Peter & Joubert, Marina. 2019. The conflation of motives of science communication: Causes, consequences, remedies. *Journal of Science Communication* 18(3). Y01.
- Wright, Nick & Nerlich, Brigitte. 2006. Use of the deficit model in a shared culture of argumentation: The case of foot and mouth science. *Public Understanding of Science* 15(3). 331–342.
- Wynne, Brian. 2006. Public engagement as a means of restoring public trust in science—hitting the notes, but missing the music? *Public Health Genomics* 9(3). 211–220.

Appendix A Distribution of survey participants by demographic attributes

Academic rank	<i>n</i>	Age	<i>n</i>
PhD candidate	124	21–30	89
Assistant/Associate professor	110	31–40	158
Full professor	101	41–50	116
Lecturer/Instructor	81	51–60	98
Research associate (post-doctoral)	48	61–70	49
Other	41	71–	30
Retired	35	Prefer not to say/NA	12
Prefer not to say/NA	12		
TOTAL	552	TOTAL	552

Gender	<i>n</i>	Region	<i>n</i>
Female	320	Europe	290
Male	193	Anglosphere	157
Other	11	Other regions	82
Prefer not to say/NA	28	Prefer not to say/NA	23
TOTAL	552	TOTAL	552

#	Country of residence	<i>n</i> ^a
1	United States	82
2	Germany	56
3	United Kingdom	48
4	Sweden	42
5	Finland	39
6	Switzerland	32
7	Nigeria	29
8	Belgium	23
9	Denmark	16
10	Netherlands	13
11	Canada	12
12	Norway	12
13	Australia	11
14	Japan	10
15	Brazil	9
16	Russia	9
17	France	8
18	Italy	7
19	Estonia	5
20	Malaysia	5
21	Czech Republic	4
22	Portugal	4
23	South Africa	4
24	Spain	4
25	People's Republic of China	3
26	Greece	3
27	Ireland	3
28	Israel	3
29	México	3
30	Austria	2
31	Chile	2
32	Croatia	2
33	Hungary	2
34	Latvia	2
35	Philippines	2
36	Poland	2
37	Serbia	2
38	Taiwan	2
...
TOTAL		552

^a Altogether 50 different countries of residence. The table includes all countries represented by at least 2 respondents.

#	Primary affiliation	<i>n</i> ^a
1	Stockholm University	15
2	University of Helsinki	14
3	University of Berne	12
4	University of Amsterdam	10
5	University of Turku	10
6	Universität Konstanz	9
7	Universiteit Antwerpen	9
8	Uppsala University	9
9	Copenhagen University	6
10	Universität Köln	6
11	Danish Language Council	5
12	Friedrich-Schiller Universität Jena	5
13	Linnaeus University	5
14	Ohio State University	5
15	University of Arizona	5
16	University of Bergen	5
17	University of Oulu	5
18	University of Tartu	5
19	Örebro University	4
20	Universidade de Lisboa	4
21	University of Fribourg	4
22	University of Tromsø	4
23	Ghent University	3
24	None	3
25	Shanghai Jiaotong University	3
26	Tomsk State University	3
27	Univ. Nacional Autónoma de México	3
28	University of California, Los Angeles	3
29	University of Ibadan	3
30	University of Joensuu	3
31	University of Melbourne	3
32	University of Queensland	3
33	University of York	3
...
TOTAL		552

^a Altogether 247 different universities/institutes. The table includes all universities/institutes represented by at least 3 respondents.

#	Linguistic subfield (OLAC)	<i>n</i> ^a
1	Sociolinguistics	173
2	Applied linguistics	162
3	Language acquisition	120
4	Discourse analysis	118
5	Pragmatics	115
6	Semantics	89
7	Syntax	81
8	Text and corpus linguistics	81
9	General linguistics	78
10	Other	76
11	Psycholinguistics	72
12	Phonology	64
13	Historical linguistics	60
14	Anthropological linguistics	55
15	Morphology	54
16	Phonetics	54
17	Cognitive science	51
18	Linguistic theories	50
19	Typology	45
20	Computational linguistics	41
21	Translating and interpreting	38
22	Language documentation	36
23	Linguistics and literature	25
24	Philosophy of language	19
25	Lexicography	17
26	Writing systems	17
27	Neurolinguistics	8
28	History of linguistics	7
29	Mathematical linguistics	7
30	Forensic linguistics	5

^a Respondents indicated between 0 and 16 subfield affiliations (median = 3).

Linguistic macro-field ^a	<i>n</i>
Function	28
Grammar	27
Applied	22
Cognition	21
General	15

^a Categorization used in the regression models

Appendix B Categorization of the linguistic subfields into macro-fields

Applied	Cognition	Function
Applied linguistics	Psycholinguistics	Anthropological linguistics
Translating and interpreting	Neurolinguistics	Discourse analysis
Lexicography	Cognitive science	Pragmatics
Linguistics and literature	Language acquisition	Sociolinguistics
Forensic linguistics	Computational linguistics	
Writing systems		
General	Grammar	
General linguistics	Phonetics	
Typology	Phonology	
Language documentation	Morphology	
Linguistic theories	Semantics	
History of linguistics	Syntax	
Mathematical linguistics	Text and corpus linguistics	
Philosophy of language	Historical linguistics	

Contact information:

Tomas Lehecka
 Faculty of Arts, Psychology and Theology
 Fabriksgatan 2
 20500 Åbo
 Finland
 e-mail: tomas(dot)lehecka(at)abo(dot)fi

Jan-Ola Östman
 Scandinavian Languages
 University of Helsinki
 e-mail: jan-ola(dot)ostman(at)helsinki(dot)fi