# Native speakers in A level modern foreign languages

# Contents

# Authors

This report was written by Rachel Taylor and Nadir Zanini from Ofqual's Strategy, Risk and Research directorate.

# Acknowledgements

The authors gratefully acknowledge the support of the students and teachers who participated in this research. The authors are also grateful to the exam boards for providing the entry and results data to facilitate the project.

# 1 Executive summary

In summer 2016, Ofqual conducted research to consider the presence of native speakers in A level modern foreign languages (MFL). The research was conducted in response to stakeholder concerns that the proportion of native speakers sitting A level MFL is increasing, and that, as a result, students are being disadvantaged due to the methods via which standards are maintained. Such concerns relate to the grades that both native and non-native speakers achieve: native speakers are considered to not be appropriately rewarded for their performance, while non-native speakers are considered to be disadvantaged due to the presence of native speakers.

The research focused on five A level MFLs – French, German, Spanish, Italian and Russian. It sought to answer 3 research questions: i) what is the percentage of native speakers sitting each A level MFL; ii) how do native speakers perform on the A level assessments compared to non-native speakers, and iii) what are the potential implications for maintaining standards. Due to the low response rate (less than 28% of schools responded to the survey) and the smaller entries in Italian and Russian, this report focuses on French, German and Spanish.

There are 4 key findings:

- As anticipated, identifying which students in the A level cohort are native speakers is challenging. This has implications for the other findings in this paper and any further research or action that might be taken based on this research.
- The percentage of native speakers (according to teacher reports) differs across subjects and is generally small but not insignificant. When considering all students in our sample, the percentage of native speakers was 8.7% in French, 17.4% in German and 10.1% in Spanish. The percentage of native speakers in our sample was lower when considering just those students that are matched to their prior attainment (ie the students that are included in the statistical predictions used to guide the setting of grade boundaries) – 7.6% in French, 10.5% in German and 4.9% in Spanish. This suggests that any effect of native speakers on the maintenance of standards might be smaller than anticipated by stakeholders.
- Native speakers outperform non-native speakers on the overall A level qualification when controlling for their prior attainment, gender and school type. The size of the effect is greatest in A level German and differs depending on students' prior attainment: native speakers outperform non-native speakers to the greatest extent for the students with lower prior attainment.
- Native speakers outperform a prior attainment based prediction at grade A by a considerable extent, particularly in A level German. This suggests that the

value added from mean GCSE to A level for native speakers is different to non-native speakers. However, due to the low percentage of native speakers sitting A level French and Spanish that are matched to their prior attainment the effect on the overall outcomes is relatively small, yet larger in German. Moreover, given the low number of native speakers, these analyses must be treated with caution.

Overall, these findings suggest that native and non-native speakers perform differently on the A level assessment and that there could be implications for the maintenance of standards. However, quantifying the size of any effect is challenging due to the low response rate in our study and the challenges of identifying native speakers. Furthermore, any implications for the maintenance of standards depend on the extent to which the proportion of native speakers in each A level MFL cohort has changed over time. If the proportion of native speakers has remained stable, there are unlikely to have been implications for maintaining standards, assuming that native speakers themselves perform similarly each examination series. While determining whether there have been any changes in the proportion of native speakers taking each A level MFL is not possible from this research alone, contextual information suggests that the proportion of native speakers taking these A level MFLs is likely to have increased over time.

On balance, this research therefore suggests that there is likely to be a small, yet important effect, of native speakers in A level MFL. However, given the challenges of identifying native speakers, the research also suggests that routinely monitoring the presence of native speakers in A level MFL each year would not be possible, and attempts to do so would not be proportionate. It is therefore recommended that thought is given to whether an adjustment to the standards is appropriate.

# 2 Introduction

In summer 2016, Ofqual conducted research to explore the presence and potential impact of native speakers on the maintenance of standards in A level modern foreign languages (MFLs). The research was conducted in response to stakeholder concerns that the proportion of native speakers sitting A level MFL is increasing, and that, as a result, students are being disadvantaged due to the methods via which standards are maintained. The research focuses specifically on the potential effect of native speakers, although it is noted that there are wider concern amongst stakeholders around the perceived severity of the grading in A level MFL.

This report discusses the findings from this research. The following section provides an overview of recent concerns around the grading of A level MFL in relation to the perceived impact of native speakers, and places these concerns in context by considering the methods via which standards are set and maintained at A level.

## 2.1    Background

A level qualifications in modern foreign languages (MFL) are available in a number of subjects, the most popular being French, German and Spanish. These qualifications were originally conceived as second language qualifications aimed at students learning the subject as an additional or foreign language. As such, the qualifications aim to develop the necessary skills to communicate and interact with users of the language being studied, and foster the ability to learn an additional language (DfE, 2015).

The A level MFLs offered by the exam boards have no formal pre-requisites for studying each language, other than the recommendation that students possess the skills and knowledge equivalent to those required at GCSE (eg see AQA, 2013)[1]. Furthermore, there are no limits on who can enter the qualifications. A level MFLs are therefore taken by students learning the subject as an additional or foreign language, as well as students who have some experience and exposure to the language outside of studying it at A level. Indeed, many educators (eg see Baker & Eversley, 2000) argue that fostering languages spoken at home is an important resource, not only for children and their families, but also for society as a whole.

The entries for A level MFLs are modest in comparison to some A level subjects and have generally been in decline over recent years (Board & Tinsley, 2016). Figures published by the Joint Council for Qualifications (JCQ) show that entries for both A level French and German decreased by around a third between 2007 and 2016, resulting in around 7,000 fewer students studying these subjects (JCQ, 2016; see

---

[1] Note that individual schools and colleges (referred to as 'schools' throughout) might have their own entry requirements.

Figure 1). While the opposite trend was observed in A level Spanish and other modern languages, this rise does not compensate for the decline in French and German. Consequently, fewer students are studying a language at A level, leading to a decline in the number of students progressing to a language degree (HESA, 2016; 2017).
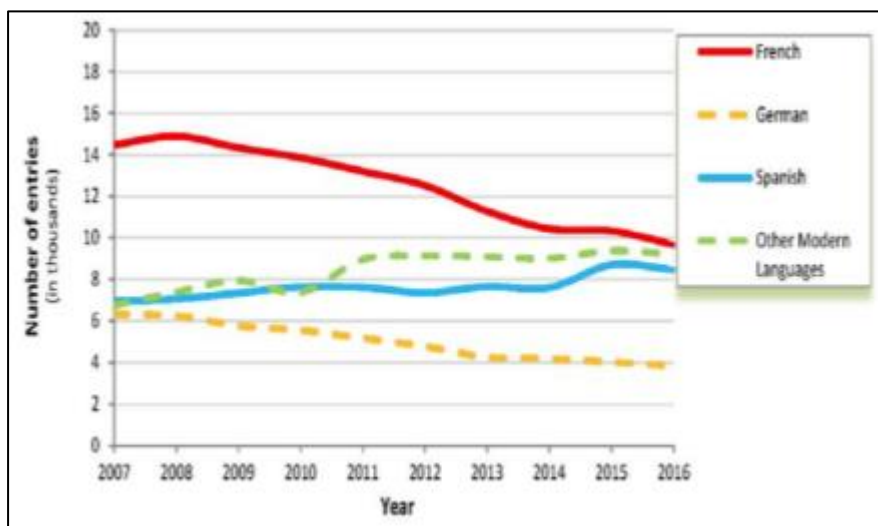


Figure 1. *A level MFL entries 2007-2016 (reproduced from JCQ, 2016, pg 1)*

The decline in the uptake of A level MFL has been linked to a number of factors. Foreign languages are often considered to be comparatively more difficult than other A levels, as highlighted in quantitative investigations based on Rasch analysis (Coe, 2008) and qualitative studies of examination materials (Smith et al., 2012). More recently, empirical research has shown that students who take French, German and Spanish as part of their A levels typically achieve their lowest grade in these subjects (Vidal Rodeiro, 2017). While there are a number of possible explanations for these findings (eg differences in student motivation, teaching etc), there are concerns that the grading in A level MFL is severe and that students are disinclined to study a language because of this. Such concerns are most frequently associated with the larger entry MFLs (ie French, German and Spanish).

Stakeholder concerns around the grading of A level MFL are complex (see Ofqual, 2014), but one of the key issues relates to the perceived impact of native speakers (see Black, 2014)[2]. There are concerns that the proportion of native speakers sitting A level MFLs are increasing as the overall entries decline, and that, as a result, there is an adverse effect on outcomes due to the methods via which standards are

---

[2] While the term 'native speaker' has been controversial in itself (Rampton,1990), it is used throughout this paper since it is in keeping with the current debate around A level MFL grading.

maintained[3]. More specifically, that increases in the proportion of native speakers is distorting the grading. This is considered to impact on the grades that native speakers achieve (ie they are not appropriately rewarded for their performance), and the grades that non-native speakers achieve (ie they are less likely to achieve the top grades due to the presence of native speakers). While many of these concerns are anecdotal, they were also reported in a recent review of the grading and take-up of A level MFL (see Ipsos Mori, 2014). Here, several students and teachers cited examples of native speaker students not attaining the top grades, even when they were fluent in the language.

Evaluating concerns from stakeholders in relation to the effects of native speakers in A level MFL is challenging, since there is little evidence that either quantifies the presence of native speakers or considers how they perform on the assessments[4]. Indeed, only contextual evidence that provides some indication of the likely numbers of native speakers sitting A level MFLs and how this might have changed is available. One potentially useful insight is offered by data from the National Pupil Database (NPD) held by the Department for Education (DfE). The figures reported in Figure 2 (panel a) show that in 2015 the percentage of students for whom English is an additional language (EAL)[5] is greater in A level French, German and Spanish than in ten of the most popular A level subjects. Furthermore, while there has been an overall reduction in EAL students taking the most popular A level subjects from 2014 to 2015, the percentage of EAL students taking MFL at A level has increased by around 2%. Although this data suggests that the proportion of EAL students taking A levels in MFL is higher than in other subjects, they do not provide information on the language these students are native/fluent in.

A second source of contextual evidence are figures published by the Office of National Statistics each year[6]. This data estimates the number of individuals within the UK population by country of birth (Figure 2, panel b). While this data suggests that the number of individuals in the UK that were born in France, Germany and Spanish has increased in recent years, it does not distinguish between school age individuals and the rest of the population, so must be interpreted with some caution.

---

[3] Concerns have also been raised that the presence of native speakers distorts the assessment since it influences expectations (see Black, 2014), though this issue is not discussed in this paper.
[4] A body of research exists considering differences in performance between monolinguals and bilinguals. However, the differences in context mean that the extent to which any findings are relevant to the present research questions is debatable.
[5] For further details on this information, see Section 3.1.
[6] At the time of writing the last release of official data on immigration was from 2015 (https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/international migration/datasets/populationoftheunitedkingdombycountryofbirthandnationality).

*a. EAL pupils taking A levels, 2014-2015*
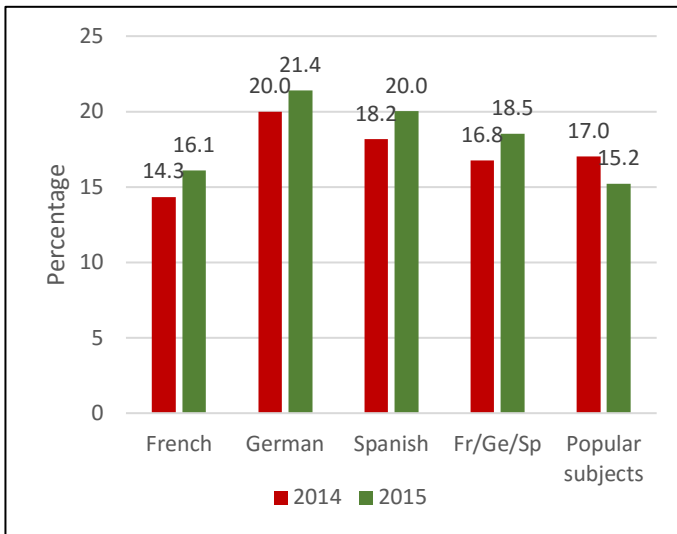
*b. UK population by country of birth, 2004-2015*



Source: NPD, 2014, 2015
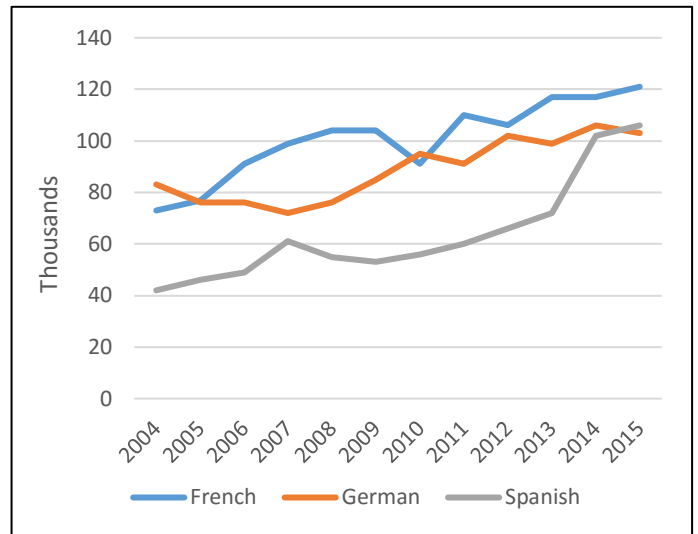
Source: ONS, 2004-2015

Figure 2. *Contextual evidence on the presence of native speakers taking A level MFL*

Given the paucity of empirical evidence and growing stakeholder concerns, this research sought to consider the presence of native speakers in A level MFL and the potential effects on standard setting. Prior to considering the research in detail, the following section places stakeholder concerns in context by considering the methods via which standards are set and maintained at A level.

## 2.2 Setting and maintaining standards in A level qualifications

At A level, standards are set and maintained using a combination of statistical and judgemental evidence, an approach that has been described in the literature as weak criterion referencing (Baird, Cresswell & Newton, 2000) or, more recently, attainment referencing (Newton, 2011). Since 2010, maintaining standards at A level has been guided by the comparable outcomes approach (Ofqual, 2015). The premise of this approach is that if the cohort of students sitting a qualification one year are comparable to the students sitting the same qualification in a previous year, then the outcomes should be similar. In practice, this is achieved by using prior attainment based predictions[7] and judgemental evidence from senior examiners.

Prior attainment based predictions model the relationship between prior attainment and outcomes in a reference year (a previous examination series), then use this relationship to predict outcomes for the current cohort of students (given their prior attainment)[8]. For A levels, the predictions use mean GCSE grade as a measure of prior attainment to predict A level grades[9]. The statistical predictions are generated at the cohort level for a specific group of students, typically those who would be expected to certificate in that qualification, in the case of A levels, 18-year-olds. Since the predictions model the relationship between prior attainment and outcomes they only include those students that are matched to their prior attainment (known as 'matched' students). Predictions are only used to guide the setting of grade boundaries where there are a sufficient number of 'matched' students – typically 500 students (Ofqual, 2016a). Where the numbers are lower, the predictions are likely to be less reliable and alternative evidence is used instead.

The prior attainment based predictions are a value added model, meaning that they model the expected value added from GCSE to A level. In relation to A level MFL, there are concerns from stakeholders that the value added relationship from GCSE to A level is different for native and non-native speakers when taking an A level MFL. More specifically, that native speakers perform differently on the assessment to non-native speakers – ie they perform better than their prior attainment would suggest – meaning that the statistical predictions under-predict outcomes. This is considered to have an adverse effect on the grades that both native and non-native speakers achieve.

While the perceived impact of native speakers on the statistical predictions might seem straightforward on the surface, there are a number of issues to consider. First, for native speakers to impact on the maintenance of standards they must perform

---

[7] Note that exam boards used statistical predictions prior to 2010 to counter possible biases in examiner judgement.
[8] See Appendix A for further details of how prior attainment based predictions are calculated.
[9] A mean GCSE score is calculated for 18-year-old students who sat at least three GCSEs 2 years earlier (at age 16).

differently to non-native speakers on the assessment, meaning that their value added from mean GCSE is different and the statistical predictions do not appropriately predict their performance. Second, there needs to be a sufficient number of native speakers sitting the assessment for this to have an effect, and these students would need to be matched to their prior attainment at GCSE. As such, any students that only sat their A levels in England (but not their GCSEs) could not influence the setting of grade boundaries, since they would not be included in the statistical predictions. Similarly, native speakers that are not 18 years old when sitting their A levels would not be considered in the statistical predictions. Third, the number of native speakers would need to have changed between the current series and the reference series (upon which the statistical predictions are based). If the number of native speakers has remained stable there should be no effect on the grading, assuming that the native speakers performed similarly in the reference series and the current series. These are all important considerations to bear in mind when interpreting any findings of this research: even if native and non-native speakers perform differently on the assessment, this does not necessarily mean that there is an issue for the maintenance of standards.

Two final points to consider are that while the statistical predictions guide the setting of grade boundaries, tolerances are applied to the differences between the outcomes and the predictions[10] and there is also a role for senior examiners' judgement. As such, outcomes can deviate from the statistical predictions within the tolerances, or awarding bodies can report outcomes that exceed the tolerances to Ofqual with supporting evidence (Ofqual, 2016b). This suggests that if there was an impact of native speakers on maintaining standards then the qualitative judgement of the examiners would not align with the statistical predictions – ie the senior examiners would find that the quality of students' work at the grade boundaries exceeded their expectations for that grade when comparing the work to scripts from the previous year – potentially to the extent that the outcomes exceed the reporting tolerances.

In recent years, this does not seem to have been the case, since none of the main A level MFL subjects (French, German or Spanish) deviated from the statistical predictions in summer 2015 or 2016 such that they were outside of the tolerances that are used by Ofqual to review outcomes (see Ofqual, 2016b). This might be because any changes are relatively small from one year to the next and therefore difficult to detect judgementally (see Ofqual, 2015). Furthermore, it might be because any effects are smaller than the reporting tolerances themselves. Indeed, it is worth noting that the only MFL subjects with reporting tolerances in summer 2016 were French, German, Spanish, and Chinese, and for many of these specifications, the

---

[10] The tolerances depend on the entry size. Where there are more than 3001 matched students the tolerances are 3%, for between 1001 and 3000 matched students the tolerance are 2%, for between 500 and 1000 matched students the tolerances are 1%, and below 500 matched students there are no tolerances.

tolerances were 2-3% (see Ofqual, 2016c). As such, there is scope for the outcomes in these subjects to deviate from the predictions to an extent, while in other MFLs, there are no reporting tolerances at all.

# 3 Research overview and aims

This research aimed to explore the potential effects of native speakers in A level MFL in summer 2016. In addition to providing evidence on the presence and performance of native speakers, the research supports Ofqual's regulatory objective to ensure that grade standards in GCSEs, AS and A levels are maintained (Ofqual, 2016d). There were 3 main aims to the study:

- to estimate the number of native speakers sitting each A level MFL;
- to compare the performance of native and non-native speakers; and
- to consider the potential effect of native speakers on the maintenance of standards.

The research focused on five languages: French, German, Spanish, Italian and Russian. The first 3 languages were included since they are the largest entry MFL subjects[11] and have most frequently been associated with concerns from stakeholders about the influence of native speakers. The remaining subjects were included since, although the entries are smaller, it was anticipated that there would be a greater proportion of native speakers sitting these subjects. Thus, they might provide additional insight into the potential effects of being a native speaker on performance.

Consideration was given to extending the research to further subjects, but any findings were thought to be less informative due to the methods via which grade boundaries are set in smaller entry subjects. Many of the concerns around the potential effects of native speakers relate to the perceived impact on the statistical predictions used to support the setting of grade boundaries. However, while statistical predictions are used to guide the process of setting grade boundaries where there are a sufficient number of students for the statistical evidence to be reliable, generally more than 500 students, in smaller entry specifications (eg the majority of other languages), statistical predictions are likely to be less reliable and alternative judgemental evidence is used instead. Thus, even if there were a large number of native speakers sitting other A level languages and they were found to perform differently to non-native speakers, it is less likely that this would influence the setting of grade boundaries.

Prior to considering the methodology for this research, the following section considers one of the key challenges: identifying who is a native speaker.

---

[11] In summer 2016, A level French, German and Spanish accounted for over 70% of the overall A level entries to MFLs.

## 3.1   Identifying native speakers

The aims of this research rely on being able to identify which students in the A level cohorts are native and non-native speakers. Whilst this might be obvious at the extremes – a student that was born and lived in France until age 16 would likely be a native French speaker, and a student who had no exposure to French outside of their A level would likely be a non-native French speaker – there is likely to be a grey area between. This could include students who might not conventionally be thought of as native speakers, but who have had significant exposure to the language they were studying beyond their experience at A level. For example, students that spend a significant amount of time with a French family each year, but do not necessarily live in a French speaking family, or students that are from a French background or family, but have lived in the UK for a long period of time. Essentially, native speakers are likely to exist on a continuum, yet for the purposes of this research, one must identify them dichotomously.

Identifying native speakers in any context is far from straightforward (Davies, 2003; Rampton, 1990), yet the definition is key to this research. If native speakers are under-identified there is a risk of concluding that there are few native speakers sitting A level MFL and that there is unlikely to be any implications for the grading. Conversely, if native speakers are over-identified, there is a risk of concluding that there is an effect of native speakers when there is none, or that there is a greater effect than there actually is. Both of these situations are undesirable since any subsequent action or lack of action might be unfounded.

As a starting point for identifying native speakers, a standard dictionary might be considered. The Oxford English Dictionary (2016) defines a native speaker as,

> *'a person for whom a specified language is their first language or the one which they normally and naturally speak, esp. a person who has spoken the language since earliest childhood, as opposed to a person who has learnt it as a second or subsequent language'*

While this definition might seem to offer a straightforward way to identify native speakers, further consideration of some of the key terms suggests that this is not the case. As discussed by Davies (2003), even terms such as 'first language' and 'second language' can be problematic: they do not necessarily encapsulate the multilingual society that individuals live in, an individual's 'first' language or the language that is naturally spoken can change over time, and an individual can (arguably) have multiple 'first' languages. These issues (and more) also apply to other terms that might be used in definitions of a native speaker such as 'mother tongue', 'dominant language', or 'home language' (Davies, 2003), suggesting that relying on a dictionary definition to identify native speakers is unlikely to be sufficient.

An alternative method would be to rely on data already held elsewhere. The National Pupil Database (NPD) contains information on students in schools and colleges in

England: this includes exam results and various student characteristics such as the first language spoken, a variable that might be used to identify native speakers. However, closer inspection of the NPD suggests that the information available is limited, since it only refers to an individual's first language. While it would therefore be possible to identify students whose first language was the language that they are studying at A level, for students with multiple languages that they are fluent in (ie a student living in a family that spoke both English and French), this would not be possible. Furthermore, the data was only complete for around 60% of the overall A level student population in June 2015 (the last year for which the data is currently available). This suggests that using the NPD to identify native speakers is unlikely to be feasible.

An alternative method is to follow the approach used in other research contexts. Perhaps the most relevant context is the study of bilingualism, where a number of studies have sought to compare the performance of monolinguals and bilinguals, thus relying on being able to identify participants according to their language status. However, like being a native speaker, bilingualism is not a clearly defined categorical variable, and the distinction between monolingualism and bilingualism has been described as 'fuzzy' at best (Luk & Bialystok, 2013). This has been problematic in bilingual research since it is not clear the extent to which bilingualism has been defined consistently in published research (Luk & Bialystok, 2013), leading to inconsistences in findings (Marian, Blumenfeld & Kaushanskaya, 2007).

While this might not seem very promising for our research, more recently a questionnaire tool to support the identification of bilingualism has been developed. This questionnaire, termed the language expertise and proficiency questionnaire (or LEAP-Q; Marian, Blumenfeld & Kaushanskaya, 2007), gathers information on all the languages that students know and use. This includes general information such as the languages that an individual knows in order of acquisition and dominance, current and average exposure to each language, and the proportion of time an individual would choose to read or speak in each language. It also includes more detailed information for each language such as age of acquisition, exposure to the language in a school/family/country where the language is spoken, language proficiency, the factors that contributed to learning the language, and current exposure to the language.

The LEAP-Q has been trialled and validated with bilingual individuals, suggesting that it is an effective tool for supporting the assessment of bilingual language status (Marian, Blumenfeld & Kaushanskaya, 2007). While it was developed for this context, many of the facets of bilingualism (and therefore the information gathered via the questionnaire) are pertinent to identifying native speakers. Indeed, any native speakers taking A level MFL will necessarily be bilinguals (if not multilinguals), since they are studying an 'additional' language in an English speaking school or college. The LEAP-Q therefore seems to offer the most promising method for identifying

native speakers in the current research, and was used to contribute to the data collection in this study.

# 4 Methodology

## 4.1 Data collection

Three sets of data were collected: entry and results data were collected from the UK exam boards that offer each of the 5 A level MFLs that were included in the research, and information on students' native speaker characteristics were collected from schools/colleges[12] and students.

### 4.1.1 Entry data

Following a request from Ofqual, each of the 5 UK exam boards (AQA, CCEA, OCR, Pearson and WJEC) provided a list of all students that were entered for each of the 5 A level language qualifications included in the research. This information allowed schools with entries to any of these A level qualifications to be contacted to gather native speaker information (see section 4.1.2).

The entry data was provided by exam boards (and was therefore complete) as of the 19 April 2016. As such, any late entries or entries that had not been processed by this date were excluded. While this meant that some schools (and students) were not included in the research, comparisons of the total number of entries (as of 19 April 2016) and the final number of certificating students showed that these figures differed by less than 3.35% (of the total entry) for each subject. This suggests that the entry data was sufficiently complete for the purposes of our research[13].

The entry data provided by the exam boards contained information relating to every student that had entered any of the 5 A level MFL subjects. This comprised: the awarding body and subject/specification being taken, the school and school type where the student studied, and a number of student level variables such as gender and date of birth. While it was therefore possible to identify which exam board each student had entered their A level MFL with, the purpose of the research was to consider any overall effects of native speakers, meaning that the data was combined across exam boards henceforth.

In addition to the variables outlined above, exam boards also identified students that were matched to their prior attainment (ie students that were included in the statistical predictions) and reported their corresponding mean GCSE score. The mean GCSE score was calculated using the same method that the exam boards use when calculating each student's prior attainment for generating statistical predictions:

---

[12] Referred to as 'schools' throughout.
[13] While it would have been possible to collect the entry data on a later date, this would have narrowed the window to gather native speaker information since many A level students go on study leave in May.

GCSE grades were converted to numbers such that A*= 8, A = 7, B = 6 and so forth, then the average grade for each student was calculated.

### 4.1.2 Native speaker information

All schools with entries for one (or more) of A level French, German, Spanish, Italian and Russian were contacted to gather native speaker information. Due to the perceived difficulties in identifying native speakers (see section 3.1), 2 sources of information were sought: teachers' perceptions of whether each student was a native speaker in the language that they were studying at A level[14], and information relating to native speaker characteristics from each student using an adapted version of the LEAP-Q[15] (see section 4.2.1). The perceptions of teachers were sought since, aside from asking the students, they are likely to be in the best position to assess students' language experience and expertise: they have regular contact with the students, and are language experts themselves. Gathering multiple sources of information was intended to increase the response rate if one piece of information was not available, and allow each measure to be validated against the other.

Each school was sent hard copies of the research materials in May 2016, with a follow-up email sent in June 2016[16]. This included: a letter outlining the purpose of the research; a set of instructions; a template for teachers to identify native speakers; a list of students taking each subject each with a unique code; a copy of the questionnaire for each student (see Appendix B); and a freepost envelope for returning the information. In total, 2,379 schools were contacted.

When contacting schools, care was taken to reassure schools and students that the research and any subsequent action was not intended to disadvantage (or advantage) one particular group of students over another. However, it is possible that some schools (or students) might have been reluctant to respond if they had concerns around this.

### 4.1.3 Results data

Results data was provided by the exam boards on 10 August 2016. The results data comprised each student's outcomes at qualification level (grade and total uniform

---

[14] Due to the issues outlined previously (see section 3.1), no definition of a native speaker was provided to teachers. Rather, they were asked to identify native speakers or students with native speaker characteristics to the best of their ability.

[15] To ensure students' anonymity (should the questionnaires be lost in transit), each student was allocated a unique code. Students were asked to write this on the questionnaire to allow their responses to be matched to the results data

[16] Note that this email was sent to all secondary school MFL teachers since it was not possible to target those schools included in the research.

mark scale [UMS] mark) and unit level (grade, raw mark and UMS mark)[17]. The results data was correct as of 10 August 2016, and therefore does not account for any post-results changes that may occur due to a review of marking or moderation. Such changes would be expected to be minor though (less than 1.1% of all A level French, Germans and Spanish qualification entries had a grade change following a post-results review in summer 2015; Ofqual, 2016e), suggesting that this is unlikely to be an issue.

### 4.1.4  Collation of the data

The 3 sets of data were collated using a student's unique identifier number. The questionnaire responses and the teacher's identification of whether students were a native speaker or not were then added to this dataset, although a number of issues were encountered. First, it was clear that the questionnaire responses from students contained large amounts of missing data, an issue that is considered further in the results section. Second, the information from teachers are not always complete either. Some teachers had provided information on the number of native speakers taking each A level language but did not provide any identifying information, meaning that it was not possible to identify which students in the school were native speakers. In addition, a number of schools had listed all or the majority of their students as native speakers, raising a question of whether this was correct or whether the schools had listed the students that had returned questionnaires instead. For the majority of these schools, consideration of the student questionnaires suggested that it was unlikely that every student listed was a native speaker, so these responses were excluded.

---

[17] For each of the five A level MFLs students must sit four units – 2 at AS and 2 at A2. One unit at each level assess speaking, and the other 2 assess reading, listening and writing. This structure is consistent across the four exam boards, although the maximum mark for each unit and the weightings differ slightly. The total qualification UMS mark is the same across all exam boards (400 UMS).

## 4.2   Materials

### 4.2.1   Student questionnaire

Information on students' native speaker characteristics was gathered using an adapted version of the LEAP-Q (see Appendix B), an instrument that has been developed for use in bilingual research (Marian, Blumenfeld & Kaushanskaya, 2007). The LEAP-Q contains 2 sections: part one collects general information on the languages that respondents possess and their language preferences, and part 2 collects detailed information on each of the languages identified in part one. Part one of the questionnaire is completed by all respondents, and part 2 is completed for each language that a respondent possesses, eg a speaker of English and French would complete part 2 twice, relating to English and French. Since the current research only focused on 5 languages, French, German, Spanish, Italian and Russian, it was only necessary to collect information relating to these languages. As such, all students were asked to complete part one of the questionnaire, then complete part 2 for the languages that they were studying at A level, eg a native French speaker who was studying A level Spanish would not be required to answer the questions relating to French.

Due to the parallels between the current research and studies considering the effects of bilingualism, much of the LEAP-Q was considered to be relevant to the present study. However, some minor changes were made to the questionnaire. First, a small number of items that were less relevant to the current research were removed, ie questions on the culture that students identify with, the number of years spent in formal education, and information on special educational needs. Second, a question was added asking for the country that students generally reside in (if this was different from the UK). This aimed to identify students who did not live in the UK but were studying their A level here, since these students are likely to be native speakers[18]. Third, some minor amendments were made to the wording of some of the questions. The LEAP-Q gathers information on students' exposure to each language, but for the purpose of this research, it was important that this related to students' experiences of the language outside of studying the subject at school. Some of the questions were therefore changed to reflect this.

The LEAP-Q was developed for use with healthy adult populations who have at least a high school level of literacy (Marian, Blumenfeld & Kaushanskaya, 2007), suggesting that it should be appropriate for use with A level students. However, to provide reassurance on the suitability of the questionnaire the revised version was piloted with a small number of year 13 students currently in their final year of an A level MFL. This resulted in a number of minor amendments being made to clarify

---

[18] In the research, very few students responded to this question and it could not be used to reliably identify native speakers.

some of the questions. To provide further reassurance on the suitability of the questionnaire, the final version of the questionnaire (and the other research materials) were sent to a number of language bodies and associations for comment.

# 5 Results

The results are organised into 6 sections: i) responses from students and teachers, ii) identifying native speakers, iii) estimating the percentage of native speakers, iv) comparing the performance of native and non-native speakers on the A level assessments at qualification level; v) comparing the performance of native and non-native speakers on the A level assessments at unit level, and vi) the implications of the presence of native speakers for maintaining standards.

Given that the structure of the A level qualifications are similar across exam boards and that each exam board used the same approach for setting grade boundaries in summer 2016, the data is pooled across exam boards for each set of analyses. This allows us to gain sample size without losing any relevant information.

## 5.1    Responses from students and teachers

Of the 2,379 schools that were contacted across the 5 languages, replies were received from 656. The number of respondents varied amongst schools, since some schools returned completed questionnaires for only some of their students. As such, the response rates are considered at the student level.

Tables 1 and 2 show the total number of students entered for each of the 5 A level MFLs in summer 2016 and the number of respondents when considering various pieces of information, for all students and matched students, respectively. This shows that the number of responses for each subject was relatively small, particularly when considering those students for whom we received complete information (ie a response from the teacher and student). However, the response rate we achieved is consistent with those generally observed in questionnaire studies (Alreck & Settle, 1995; Shaughnessy & Zechmeister, 1994).

The response rates were similar across subjects, though slightly higher in German and lower in Russian and Italian. This might reflect the motivation of individuals to respond. Indeed, many of the anecdotal concerns around the impact of native speakers on the grading of A level MFL focus on A level German. The response rates show that we received more responses from teachers than students, though this is not surprising given that the latter was more labour intensive (for schools) to collect.

Given the relatively low response rate, the sample size is an important issue to consider. The total entry for the 5 A level MFLs would be sufficient to perform a number of different kinds of analyses, yet our response rates are at best less than a third of the total entry. This is particularly problematic in Italian and Russian, where the number of students is small. Considering this, the analyses in this report focus on French, German and Spanish.

In addition to the size of the sample, the representativeness of our sample also requires consideration. However, the extent to which we can consider this depends on which students we are including in our analyses and therefore how we are identifying native and non-native speakers. This is considered in the following section, so the representativeness of our sample is discussed there.

Table 1.  *Numbers of students and respondents – all students*

|  | French | Spanish | German | Italian | Russian |
|---|---|---|---|---|---|
| Total entry | 9,460 | 8,241 | 3,786 | 822 | 1143 |
| Questionnaire | 1,930 | 1,497 | 800 | 114 | 125 |
| Teacher | 2,561 | 2,077 | 1,108 | 168 | 266 |
| Questionnaire + teacher | 1,431 | 1,114 | 628 | 68 | 98 |

*Note: 'Total entry' refers to students for which both entry and results data were available. 'Questionnaire' refers to students who returned the questionnaire (without accounting for missing data), and 'teacher' refers to students that the teacher identified as a native or non-native speaker.*

Table 2.  *Numbers of students and respondents – matched students*

|  | French | Spanish | German | Italian | Russian |
|---|---|---|---|---|---|
| Total matched | 7,299 | 6,005 | 2,850 | 482 | 499 |
| Questionnaire | 1,528 | 1,165 | 641 | 81 | 55 |
| Teacher | 1,991 | 1,556 | 849 | 89 | 123 |
| Questionnaire + teacher | 1,137 | 876 | 485 | 41 | 40 |

*Note: 'Matched' refers to matched to their prior attainment for which both entry and results data were available. 'Questionnaire' refers to students who returned the questionnaire (without accounting for missing data), and 'teacher' refers to students that the teacher identified as a native or non-native speaker.*

## 5.2 Identifying native speakers

To identify native speakers, 2 sets of information were collected – teachers' perceptions of whether students were a native speaker or not in the language they were studying at A level, and information relating to students' language exposure and proficiency via an adapted version of the LEAP-Q (Marian, Blumenfeld & Kaushanskaya, 2007). It was intended that both sets of information would be used to identify native speakers: the questionnaire data would be used to develop a method for identifying native speakers, then this would be validated against the teacher responses. While the LEAP-Q has been validated for use with bilingual and multilinguals, the items of the questionnaire are not routinely combined to form subscales assessing bilingual or multilingual status (Marian, Blumenfeld & Kaushanskaya, 2007). As such, several pieces of information from the questionnaire were considered as a method for identifying native speakers: i) the languages known by students; ii) when students acquired each language; and iii) exposure to the language now and when learning the language.

The number of questionnaire responses received for each language has been shown previously in Table 1. While they are sufficient to conduct a range of analyses, it is important to note that completed questionnaires were received from a much smaller proportion of students, around 20% considering the 5 languages together. As such, the analyses in this section each include students who had answered the relevant questionnaire items (indeed any method for identifying native speakers based on the questionnaire data would only include those students with the necessary information).

### 5.2.1 Known languages – by dominance and order of acquisition

Students were asked to report up to 5 languages that they know in order of dominance and this is reported by A level MFL in Table 3. As shown, only a small number of students taking each A level MFL reported the language that they were studying as their dominant language, for example only 21 students (1.4%) indicated that French was their dominant language, while 80.1% reported that French was their second dominant language, and the remaining 18.4% reported it as their third or less dominant language. This trend is similar for A level Spanish and German, though in these cases, the proportion of students reporting the language they are studying as their dominant language is slightly higher. Nonetheless, the vast majority of students did not report their A level language as their dominant language.

Table 3. *Students by order of dominance of language studied at A level*

|  | French | | German | | Spanish | |
|---|---|---|---|---|---|---|
|  | N | % | N | % | N | % |
| First | 21 | 1.4 | 17 | 2.7 | 36 | 3.1 |
| Second | 1,199 | 80.1 | 489 | 77.5 | 840 | 73.3 |
| Third | 234 | 15.6 | 114 | 18.1 | 252 | 22 |
| Fourth | 37 | 2.5 | 9 | 1.4 | 15 | 1.3 |
| Fifth | 5 | 0.3 | 2 | 0.3 | 3 | 0.3 |
| Total | 1,496 | 100 | 631 | 100 | 1,146 | 100 |

*Note: The percentages are computed by language. Percentages may not total to 100% due to rounding.*

Students were also asked to report up to 5 languages that they know in order of acquisition, with their native language first (Table 4). As for the language of dominance, the vast majority of students reported that the language studied at A level was their second (or third) language in order of acquisition. This pattern is similar across the 3 subjects, although for A level Spanish, a relatively high proportion of students reported that Spanish was the third or fourth language that they acquired. This provides some indication that for the majority of students, the language that they were studying at A level was an additional language, rather than their native language.

Table 4. *Students by order of acquisition of language studied at A level*

|  | French | | German | | Spanish | |
|---|---|---|---|---|---|---|
|  | N | % | N | % | N | % |
| First (native) | 42 | 2.8 | 28 | 4.5 | 26 | 2.3 |
| Second | 1175 | 78.8 | 375 | 59.7 | 602 | 52.6 |
| Third | 234 | 15.7 | 190 | 30.3 | 437 | 38.2 |
| Fourth | 38 | 2.5 | 32 | 5.1 | 73 | 6.4 |
| Fifth | 3 | 0.2 | 3 | 0.5 | 7 | 0.6 |
| Total | 1,492 | 100 | 628 | 100 | 1,145 | 100 |

*Note: The percentages are computed by language. Percentages may not total to 100% due to rounding.*

From the information on language acquisition, it was possible to determine whether the first language that students acquired was English, the MFL taken at A level, or another language[19]. Table 5 shows that around 90% of A level French and Spanish

[19] Note that the total number of respondents in Table 5 is slightly higher since this is only based on the first language that students acquired.

students reported that the first language they acquired was English, with only a slightly lower percentage in German (87%).

Table 5. *First language of acquisition (native language) as reported by students, by A level MFL studied*

|  | French | | German | | Spanish | |
|---|---|---|---|---|---|---|
|  | N | % | N | % | N | % |
| English | 1366 | 89.5 | 554 | 86.7 | 1049 | 90.1 |
| A level language | 42 | 1.4 | 28 | 4.4 | 26 | 2.2 |
| Other language | 118 | 7.7 | 57 | 8.9 | 89 | 7.6 |
| All (non-missing) | 1526 | 100 | 639 | 100 | 1164 | 100 |

*Note. The percentages are computed by language. Percentages may not total to 100% due to rounding.*

The information relating to students' dominant language and order of acquisition provides some indication of whether students are native speakers or not in the subject that they are studying at A level. Indeed, it would be possible to identify some students as native speakers based on this information alone, ie those that reported their A level MFL as their dominant (or native) language. Using this information would suggest that the percentage of native speakers in each subject is very small and almost negligible, apart from German (Table 5). For the majority of students, however, this information would be insufficient to reliably identify native and non-native speakers – ie for those students that have exposure to the language that is not captured here, or those that acquired the language as their second language but at a very young age. As such, using this information to identify native speakers is likely to significantly under-estimate the presence of native speakers in each subject, and it should therefore not be used (in isolation) to identify native speakers in this study.

### 5.2.2 Language acquisition

An alternative approach to identifying native speakers might be to consider the age that students began acquiring the language they are studying at A level or exposure to the language. This information can help to understand the role of family and social background on the development of the language, outside the school environment.

Figure 3 shows a distribution of the age when students began to learn their A level language (as reported by students), alongside information from the teachers that identified students as a native or non-native speaker (where this information was available). For the 3 subjects, there is a peak at 11 years of age. This is the age at which students enter key stage 3 and is therefore one of the points in their education that they are likely to start learning a foreign language.

Among students taking A level French, a relatively large number of students began acquiring the language between age 6 and 10, while for Spanish and German,

students tended to begin acquiring the language around age 12 to 13 (and age 14 for Spanish). This suggests that some students began learning a language in primary education, while others were introduced to the study of languages in secondary education, possibly to allow them to sit a GCSE in the language (this might have been influenced by the inclusion of languages in the EBacc)[20].

From Figure 3 it is also apparent that a small (but not insignificant) number of students reported acquiring the language they are taking at A level before the age of 5, suggesting a long-lasting exposure to the language that started well before school age. It would be natural to think that these students are likely to be native speakers, yet this does not always seem to be the case. While students who began acquiring the language at a very young age were mostly (independently) identified as a native speaker by their teacher, some were identified as non-native speakers. Furthermore, native-speakers (as identified by teachers) appear to be widespread across the age range, including some students who reported that they began learning the language post-16. This suggests that the age of acquisition does not correspond well to the reports from teachers. This might be because the age of acquisition was not reliably reported by students, or because the teachers did not reliably identify whether students were native speakers or not. Consideration of some of the questionnaire responses from students suggested that there were some issues in reporting the age of acquisition, since some students reported a late age of acquisition and also reported that they had significant exposure to the language in a country, family, or school where the language is spoken. This seems implausible and suggests that the age of acquisition is not sufficiently reliable to be used to identify native speakers in this research.

It is worth considering that even if the information was reliably reported, determining an age of acquisition that one could use to distinguish between native and non-native speakers is not really possible without considering additional information. For example, a student learning a language from birth might rightly be considered to be a native speaker, yet what of the student that started learning the language at age 3? They might be a native speaker or might have had some language exposure, however limited, at pre-school.

---

[20] The EBacc is a school performance measure introduced in 2010. To achieve the EBacc students must achieve a grade C or above in five GCSE subjects: English, mathematics, the science, an MFL and history or geography.
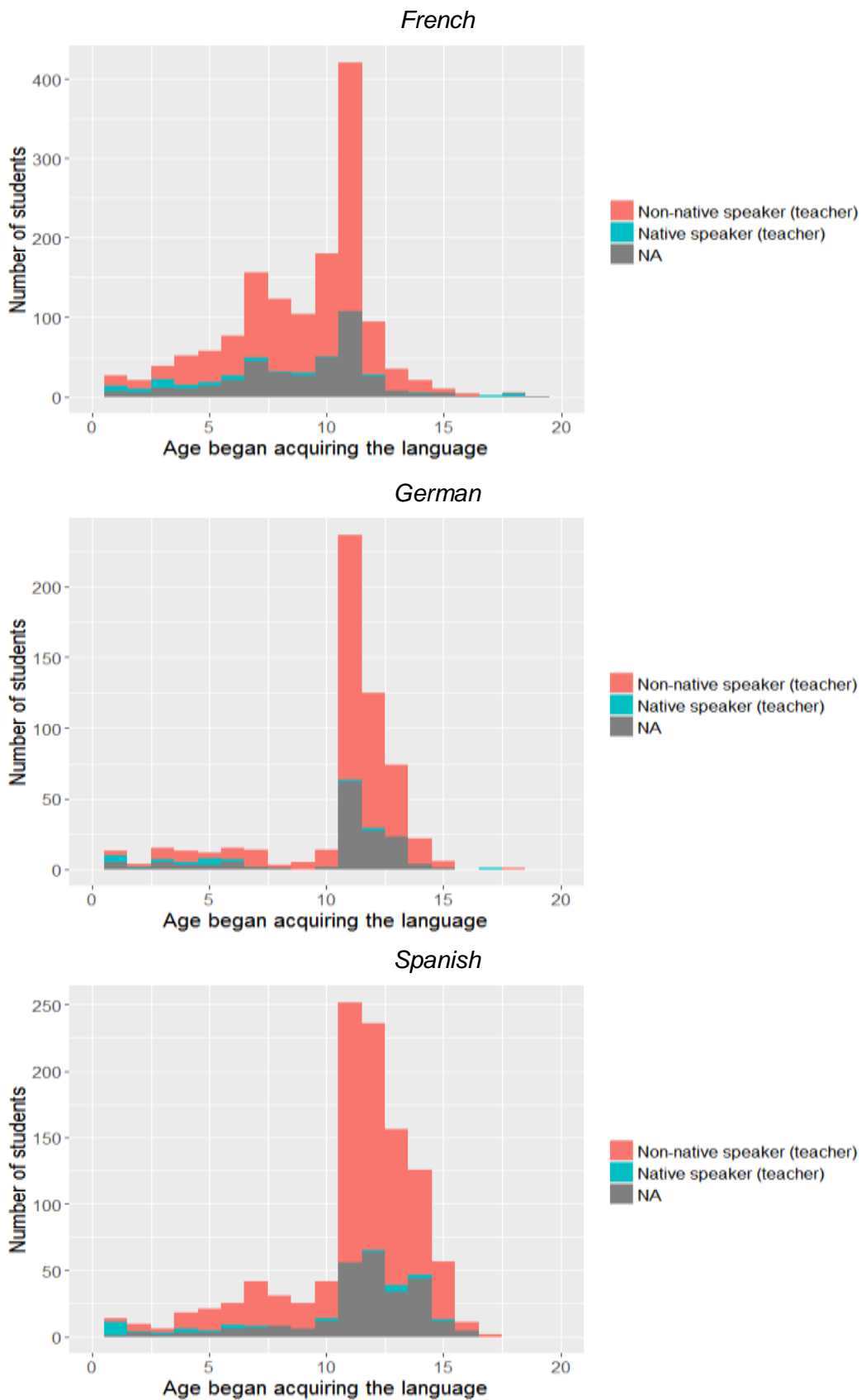
Figure 3. *Distribution of self-reported age when acquiring the language studied at A level, for native and non-native speakers (as reported by teachers)*

In addition to age of acquisition, further information on the exposure to a foreign language was collected including the extent to which students have spent time in a country, a family, or a school where the language is spoken. This data is summarised (in years) in Table 6, for students who were independently identified by their teacher as a native or non-native speaker. This shows that, in general, students identified as native speakers by their teacher have had more exposure to the language they are studying than students identified as non-native speakers.

Despite this general pattern, some of the figures suggest that there are inconsistencies between the data provided by students and teachers. As an example, one would not expect any non-native speakers in the language being studied at A level (perhaps with a few exceptions) to have lived in a family speaking this language. However, for A level French, over 6% of respondents had, with some reporting that they had lived in a French speaking family for 11 years or more (3.5%). The figures are similar, though slightly lower, for German and Spanish, and although the number of students is quite small, this is a non-negligible fraction of the total number of respondents.

These inconsistencies raise questions relating to whether or not one or both sources of data are genuine. If they are both genuine, this suggests that some students, exposed to a language for a significant period of their life are not native-speakers. This seems unlikely, suggesting that some students and/or teachers have mistakenly reported (some of) the information required.

Closer inspection of the questionnaire responses suggests that some students might have confused the question about the age they began acquiring the language with the question about the number of years they have spent learning that language, since their responses seem implausible based on their responses to the rest of the questionnaire. This suggests that information on the time spent in a country/family/school where the language is spoken is not sufficiently reliable to identify native speakers. Furthermore, even if the information had been reported reliably, many of the students did not respond to these questions and determining a point at which one could separate native and non-native speakers based on their responses would be difficult. This is perhaps obvious at the extremes, ie a student who has no exposure to the language through a country/family/school where the language is spoken is unlikely to be a native speaker, while a student who has lived in a French speaking family for 18 years is likely to be a native speakers. There is a grey area between.

An alternative explanation for the inconsistencies in the data is that some of the teachers were not in a position to reliably identify native-speakers, for example, if they were less familiar with their students' familial background or were perhaps biased by other information, such as the language skills of the student. This seems less likely since the teachers are language experts themselves and have spent the

previous 18 months teaching those students. However, this is something that must be borne in mind.

Table 6. *Years spent in language environment, by subject*

| Years | | French | | German | | Spanish | |
|---|---|---|---|---|---|---|---|
| | | N | % | N | % | N | % |
| *Country* | | | | | | | |
| 0 | Native | 31 | 32.6 | 9 | 18.8 | 12 | 30.0 |
| | Non-native | 737 | 88.1 | 320 | 90.7 | 546 | 88.3 |
| 1-10 | Native | 51 | 53.7 | 23 | 47.9 | 20 | 50.0 |
| | Non-native | 93 | 11.1 | 30 | 8.5 | 70 | 11.3 |
| 11-19 | Native | 13 | 13.7 | 16 | 33.3 | 8 | 20.0 |
| | Non-native | 7 | 0.8 | 3 | 0.8 | 2 | 0.3 |
| *Family* | | | | | | | |
| 0 | Native | 27 | 28.1 | 11 | 23.9 | 19 | 46.3 |
| | Non-native | 780 | 93.6 | 334 | 94.4 | 573 | 94.9 |
| 1-10 | Native | 12 | 12.5 | 5 | 10.9 | 3 | 7.3 |
| | Non-native | 24 | 2.9 | 6 | 1.7 | 18 | 3.0 |
| 11-19 | Native | 57 | 59.4 | 30 | 65.2 | 19 | 46.3 |
| | Non-native | 29 | 3.5 | 14 | 4.0 | 13 | 2.2 |
| *School* | | | | | | | |
| 0 | Native | 52 | 55.3 | 25 | 52.1 | 22 | 56.4 |
| | Non-native | 789 | 95.3 | 344 | 96.9 | 579 | 95.4 |
| 1-10 | Native | 35 | 34.0 | 15 | 31.3 | 12 | 30.8 |
| | Non-native | 37 | 4.5 | 10 | 2.8 | 27 | 4.4 |
| 11-19 | Native | 10 | 10.6 | 8 | 16.7 | 5 | 12.8 |
| | Non-native | 2 | 0.2 | 1 | 0.3 | 1 | 0.2 |

*Note: The percentage is computed by language for native and non-native speakers separately.*

### 5.2.3 Developing a method to identify native speakers

So far, the direct questions regarding the languages known by students or the age of acquisition/length of exposure have not enabled us to reliably identify whether a student is a native speaker in the language that they are studying at A level. The student questionnaire, however, included some indirect questions aimed at gathering data on the exposure to foreign languages and on the factors contributing to the learning of foreign languages outside of school. This included 6 questions relating to students' current exposure to the language, and 6 questions relating to the factors that contributed to an individual learning the language. The questions were the same in both contexts: 2 referred to interactions with family and friends, 3 related to the

medium with which the learning happened (reading, watching or listening), and one to the degree of self-instruction (see Appendix B).

This data is summarised for all A level French students who responded to the questionnaire in Table 7 (the findings are similar for German and Spanish so only the analyses for French are presented here). This shows that current exposure and exposure when learning the language are generally fairly similar.

Table 7. *Current exposure and factors contributing to learning French outside school and school work/homework*

| | | Not at all | | Minimally | | Moderately | | A lot | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | % | N | % | N | % | N | % |
| | | *Factors contributing to learning the language* | | | | | | | |
| Interacting with friends | (t.fr) | 587 | 39.0 | 519 | 34.4 | 275 | 18.2 | 126 | 8.4 |
| Interacting with family | (t.fa) | 769 | 50.9 | 373 | 24.7 | 212 | 14.0 | 157 | 10.4 |
| Reading | (t.r) | 169 | 11.2 | 521 | 34.5 | 574 | 38.0 | 246 | 16.3 |
| Self-instruction | (t.s) | 815 | 54.2 | 419 | 27.9 | 210 | 14.0 | 60 | 4.0 |
| Watching TV | (t.w) | 378 | 25.0 | 568 | 37.6 | 410 | 27.2 | 153 | 10.1 |
| Listening to the radio/music | (t.l) | 307 | 20.3 | 587 | 38.8 | 400 | 26.4 | 219 | 14.5 |
| | | *Current exposure* | | | | | | | |
| Interacting with friends | (e.fr) | 642 | 42.6 | 577 | 38.3 | 233 | 15.5 | 56 | 3.7 |
| Interacting with family | (e.fa) | 850 | 56.4 | 388 | 25.7 | 143 | 9.5 | 126 | 8.4 |
| Reading | (e.r) | 220 | 14.6 | 639 | 42.3 | 495 | 32.8 | 156 | 10.3 |
| Self-instruction | (e.s) | 942 | 62.5 | 382 | 25.3 | 144 | 9.5 | 40 | 2.7 |
| Watching TV | (e.w) | 433 | 28.7 | 641 | 42.5 | 344 | 22.8 | 91 | 6.0 |
| Listening to radio/music | (e.l) | 358 | 23.7 | 595 | 39.4 | 372 | 24.6 | 185 | 12.3 |

*Note: The percentage is computed by row.*

It was anticipated that considering the medium via which the language was learned or current exposure might provide a method for identifying native speakers. Figure 4 shows the correlations among these 12 indicators (using the labels for the questionnaire items shown in table 7). The figure shows the correlation coefficient to the left of the variable names, and the strength of the correlation to the right of the

variable names. This shows that for each of the 6 indicators, the current exposure to the language is most strongly correlated to its counterpart among the factors contributing to learning the language, with the strongest correlation (0.87) between the 2 indicators referring to the extent to which respondents interact with family members. However, apart from some correlations between the indicators of the medium of instruction, when considering the factors contributing to the learning of French, the different indicators do not appear to be strongly related to one another. A similar pattern arises from the questions relating to current exposure to the language.
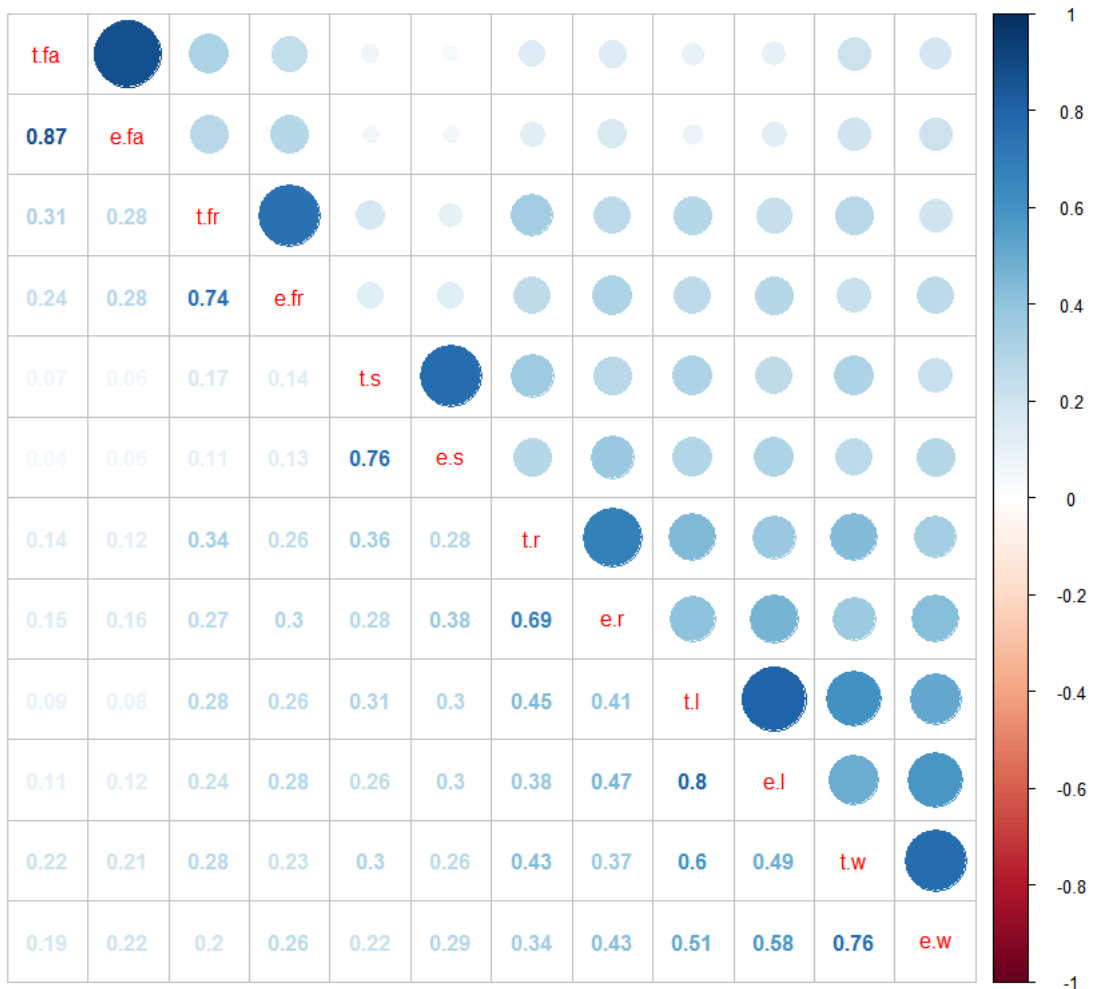


Figure 4. *Correlations between exposure and factors contributing to learning of French (for variable names see table 7)*

Despite some relatively low correlations, the feasibility of developing a method for identifying native speakers based on combining the responses to these items on the questionnaire was considered. First, for each of the 3 MFLs, a factor analysis was performed separately on *i)* the set of 6 indicators of current exposure and *ii)* the 6

factors contributing to the learning of each of the languages[21]. In both cases, the results suggested that the 6 indicators are not underpinned by the same factor: the proportion of variance explained by a single factor model was around 0.33 and Cronbach's alpha was around 0.70. Attempts to fit a 2-factor model showed better results, suggesting that indicators of exposure to the language and how the language is learnt are underpinned by 2 different factors: one relating to the interaction with family members (factor loadings around 1), and one relating to all of the other indicators (though with very low factor loadings). The proportion of cumulative variance explained by the 2-factor model appears quite low still though (just below 0.50), suggesting that a better model to fit the data should be sought.

Attempts were therefore made to combine all of the 12 indicators in one model. The results mirrored what has been presented above for the 2 sets of items separately, suggesting the presence of 2 latent factors: the first relating to the interaction with family members with respect to the learning the language (factor loading above 1) and the exposure to the language (factor loading = 0.9), and the second underlying all the other indicators. However, in this case, the proportion of cumulative variance explained by the model did not reach 50%, the standard cut-off value indicating whether a model fits the data well or not.

These analyses therefore suggest that it is not possible to develop a method for identifying native speakers by combining all the indicators relating to students' exposure to the language and the factors contributing to learning the language. However, it might be possible to identify native speakers based on the 2 indicators relating to the influence of the family on learning the language and on current exposure to the language, since these were found to be underpinned by the same latent factor. These 2 indicators were therefore combined to derive a measure of the role of family background. In order to check whether the 'family score' resulting from adding up the 2 indicators is a good indicator of whether students are native speakers or not, it is possible to contrast it with the information provided by teachers.

Figure 5 (panel a, c) shows that the fraction of native speakers with a family score of 8 is fairly high (60%)[22] and significantly greater than for any other level of family score[23]. However, there are also native-speakers spread out across the family score range. Figure 5 (panel b, d) also shows that when the family score was considered alongside the performance at A level summarised by the UMS mark, students with high exposure to the language within the family did not outperform those with low

---

[21] See Appendix C for a methodological description of the factor analysis.

[22] A score of 8 means that students reported that interacting with their family contributed to them learning the language 'a lot', and that they were currently exposed to the language via their family 'a lot'.

[23] Similar results were found for Spanish and German, though the larger fraction of native speakers with high family scores of 8 was not significantly greater than those with a family score of 7.

levels of the family score index. While they might not necessarily be expected to, overall these analyses suggests that the family score indicator does not allow native-speakers to be reliably identified.
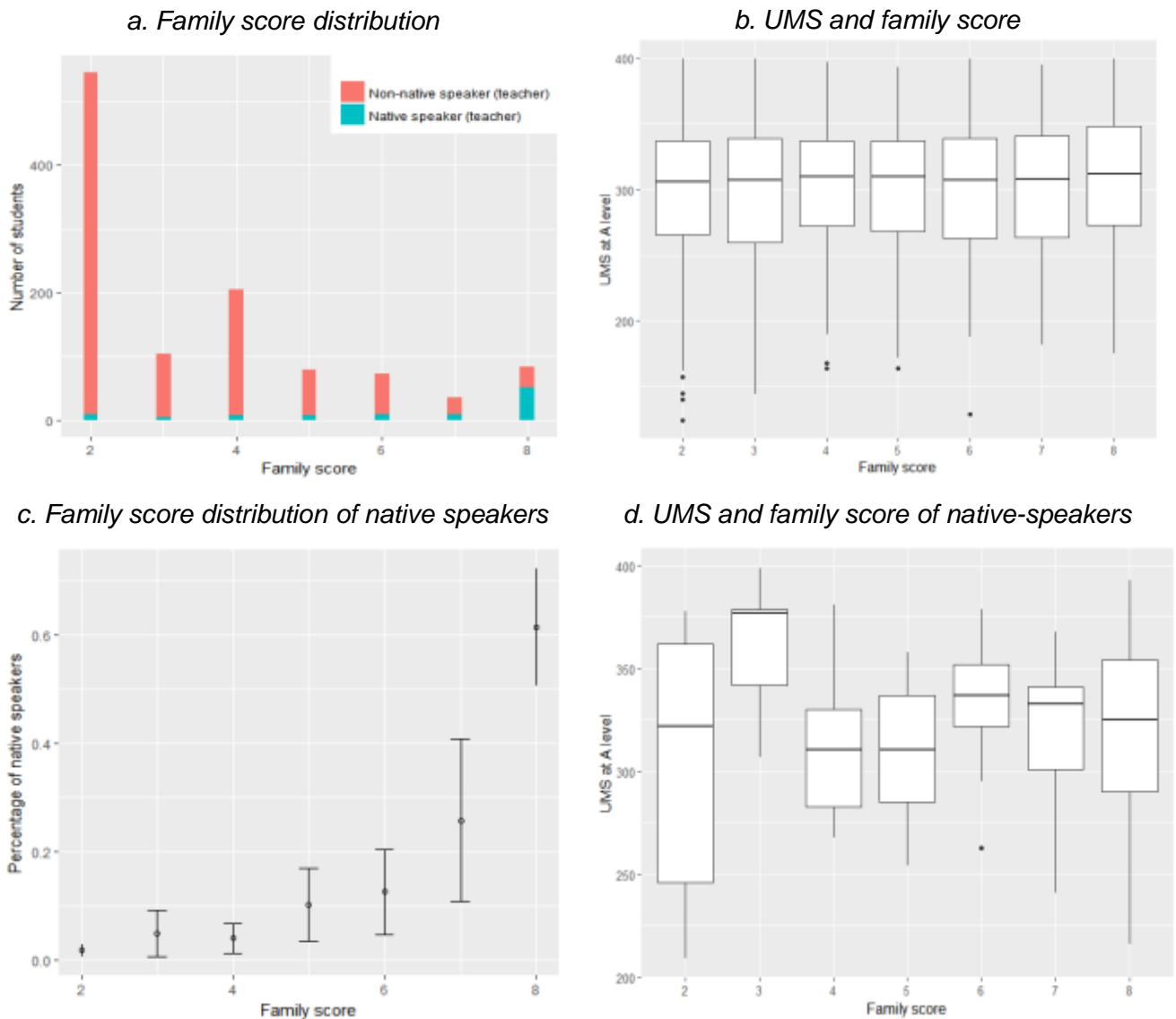


Figure 5. *The family score of A level French students*

*Note: The classification of students between native and non-native speakers is based on the information provided by teachers.*

### 5.2.4 Summary

In summary, these analyses suggest that it is not possible to use the responses to the questionnaire to reliably identify native speakers, either when considering responses to individual questions, or when combining items on the questionnaire to devise a method to identify native speakers. This might be because of issues with the data provided, eg the responses to some questionnaire items are at odds with

one another or the teachers' perceptions – or because the questionnaire did not function as intended with our respondents.

From considering the individual responses, it appears that some students mis-interpreted some of the questions. For example, when asked to report the age of language acquisition and the time spent in a county, family or school where the language is spoken, some of the responses are inconsistent or implausible: it is unlikely that a student who started learning a language at age 18 has also spent 18 years living in a family who speak that language, unless they were taking the A level as a mature student.

This confirms one of the key challenges in this research: identifying which students are native speakers. This was anticipated prior to conducting the research (see section 3.1 for a discussion), so multiple sources of information were collected to facilitate this. However, the poor functioning of the questionnaire meant that it was not possible to use the information from students to identify native speakers.

Given this, using the responses from the teachers is the only possible method of identifying native speakers. While it is anticipated that teachers will have a good understanding of whether students have native speaker characteristics in the language that they are studying given that they have regular contact with the students and are language experts themselves, there are limitations of this approach too. The sample size is limited to just those students with a response from the teacher (although this is higher than the number of students with a questionnaire response; see Tables 1 and 2), and there are some concerns around the reliability of this measure. It is possible that some teachers might not be fully aware of each students' language background, hence might under-estimate the number of native speakers. Conversely, it is possible that some teachers might think that a student is a native speaker based on their proficiency, hence might over-estimate the number of native speakers. Nonetheless, such instances should be relatively rare, and the responses from teachers are likely to be accurate for the majority of students. As such, this measure was used to identify native speakers throughout the remainder of this research

Prior to conducting the analyses, reassurance was sought around the representativeness of our final sample (ie those with a response from the teacher), compared to the overall cohort taking each A level MFL. The details of these analyses are provided in Appendix D. In general, the analyses suggest that there are only minor differences between the type of students that responded to our survey and the overall cohorts. These differences relate to the specification taken and gender, with female students more likely to reply. Our sample, though relatively small, therefore seems reasonably representative of the whole cohort of A level students in terms of mean GCSE grade, attainment in A level MFL, geographical location and type of school attended. Despite this, it is still possible that some schools were more motivated to respond to our questionnaire than others, for

example, those with a large number of native speakers, which could potentially bias our findings.

## 5.3 Estimating the percentage of native speakers

Using the responses from teachers, Tables 8 and 9 show the percentage of native and non-native speakers in our sample for each of the 3 A level MFLs, for all students and matched students, respectively. This shows that the percentage of native speakers taking each A level varies according to language, with the percentage of native speakers being greatest in A level German. The figures also vary when one considers all students or just those matched to their prior attainment (ie those students that are included in the statistical predictions), with the percentage of native speakers being lower for the latter. Given that grade boundaries are set on the basis of matched students only, this may suggest that the effect of native speakers on maintaining standards is smaller than anticipated by stakeholders.

These figures provide an indication of the percentage of native speakers in each of the A level MFL cohorts, information that has not previously been available. However, they must be interpreted with caution and can only be considered as an estimate: the figures relate to a small proportion of the MFL cohort in each subject that responded to our survey. They only relate to June 2016 entries, and they rely on our method used to identify native speakers, in this case, the responses from teachers. If an alternative method of identifying native speakers was used, or if a larger sample had been achieved, then the figures would likely vary. Furthermore, they provide no indication of how the percentage of native speakers might have changed over time, an important consideration in relation to the potential implications of native speakers for maintaining standards.

Table 8. *Native and non-native speakers – all students, by subject*

|  | French | | German | | Spanish | |
|---|---|---|---|---|---|---|
|  | N | % | N | % | N | % |
| Native speakers | 223 | 8.7 | 193 | 17.4 | 210 | 10.1 |
| Non-native speakers | 2,338 | 91.3 | 915 | 82.6 | 1,867 | 89.9 |
| All | 2,561 | 100 | 1,108 | 100 | 2,077 | 100 |

*Note. The percentages are computed by language.*

Table 9. *Native and non-native speakers – 'matched' students, by subject*

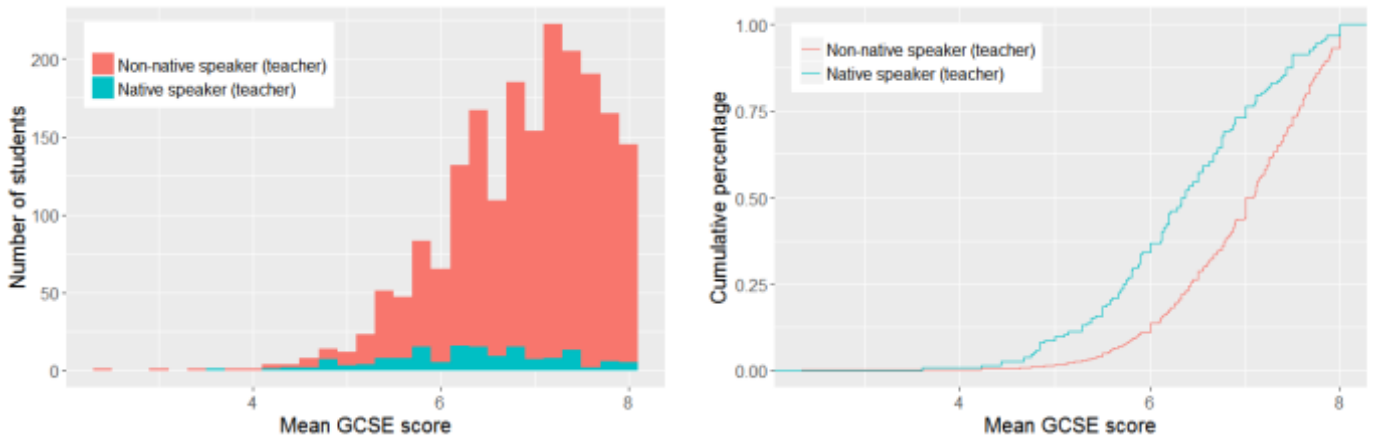|  | French | | German | | Spanish | |
|---|---|---|---|---|---|---|
|  | N | % | N | % | N | % |
| Native speakers | 152 | 7.6 | 89 | 10.5 | 76 | 4.9 |
| Non-native speakers | 1,839 | 92.4 | 760 | 89.5 | 1,480 | 95.1 |
| All | 1,991 | 100 | 849 | 100 | 1,556 | 100 |

*Note. The percentages are computed by language.*

### 5.4 Comparing the performance of native and non-native speakers at qualification level

The second research aim was to compare the performance of native and non-native speakers on the A level assessments at qualification and unit level. Initially, the native and non-native speakers were compared at qualification level on a number of basic descriptive measures to consider the composition of each group. This included the prior attainment of the students (mean GCSE score), the qualification grade achieved and the total UMS score.

The distribution of mean GCSE scores for the native and non-native speakers is show in Figure 6. This illustrates that, in general, students taking A level MFL have a relatively high level of prior attainment. However, Figure 6 also shows that although the native speakers are varied in terms of their prior attainment, the number of non-native speakers taking each language rises considerably among students with a mean GCSE score above 6 (which corresponds to an average GCSE grade B). This suggests that the nature of the native and non-native candidature is different in terms of their prior attainment.
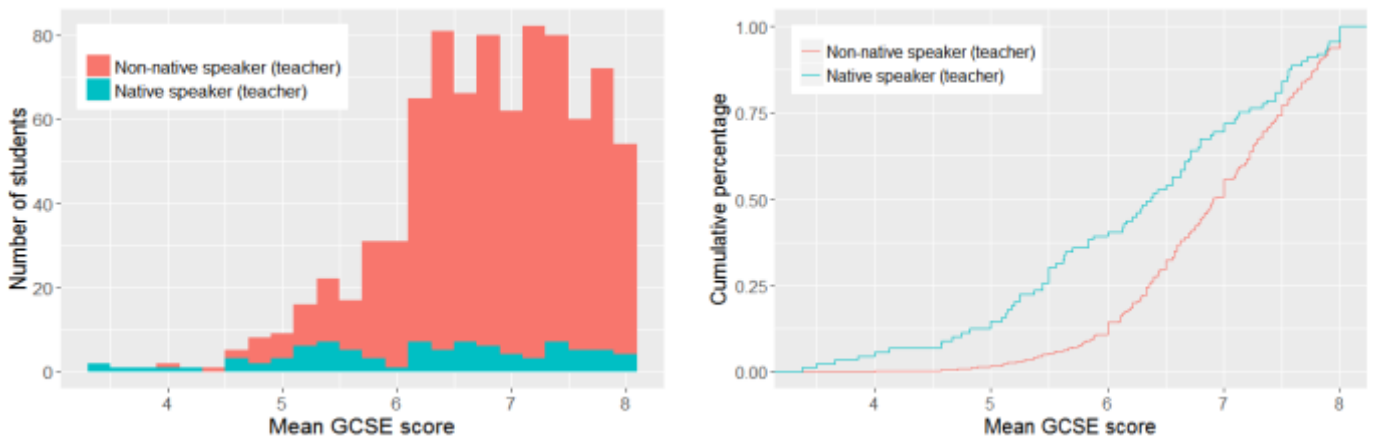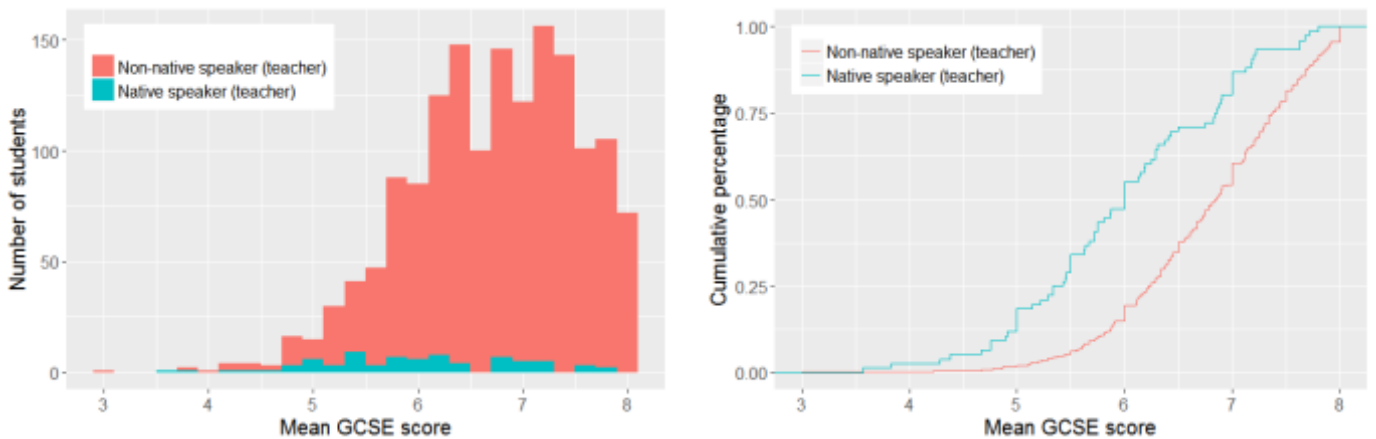
Figure 6. *Distribution (on the left) and cumulative distribution (on the right) of the mean GCSE distribution for native and non-native speakers*

A second set of descriptive statistics shows the overall A level qualification grade achieved by the native and non-native speakers (Figure 7). For French and Spanish, the proportion of native speakers achieving each grade (of the overall number of students) is relatively constant, yet for German, the situation is different: in percentage terms, there are more native speakers achieving the top grades than those achieving lower grades. This is clear at A*, where almost half of the students in our sample are native-speakers, and at grade A, where almost a fourth of the students are native speakers. For lower grades (B, C and D), the proportion of native speakers is much smaller.

*French*
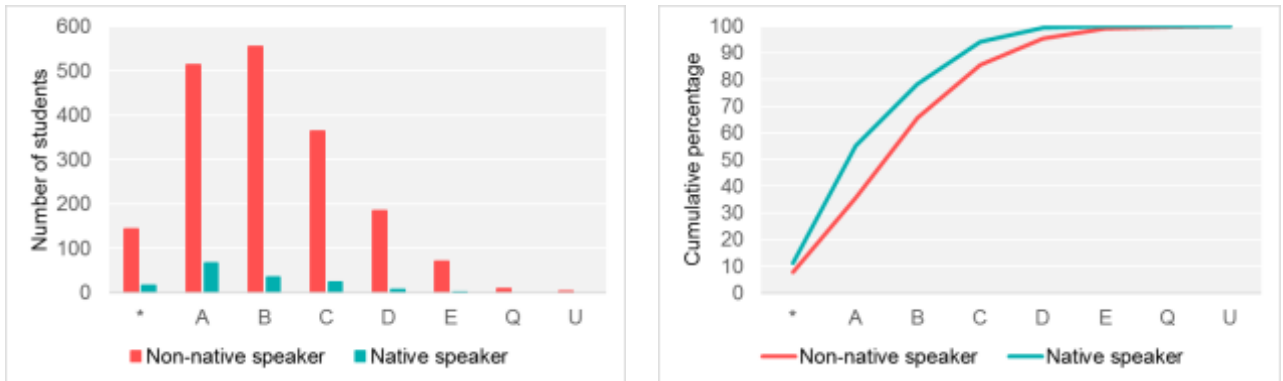


*German*
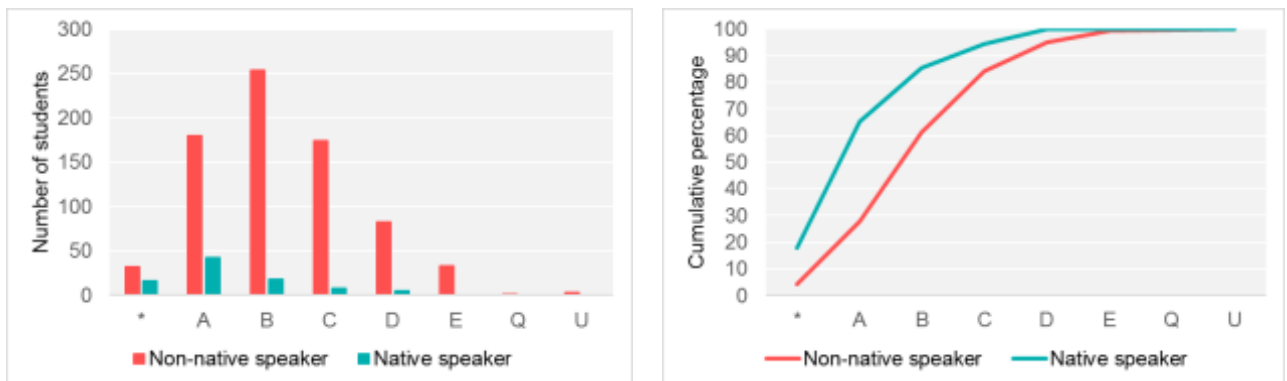


*Spanish*
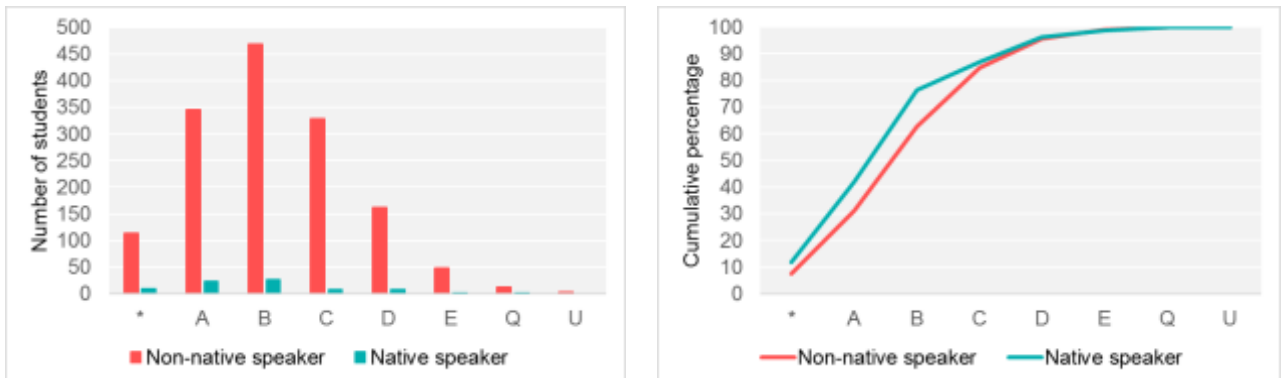


Figure 7. *Grade distributions (on the left) and cumulative distributions (on the right) for native and non-native speakers*

The third set of descriptive statistics consider the performance of native and non-native speakers in terms of UMS score, rather than grade. For each of the 3 subjects, Figure 8 illustrates the relationship between UMS score at A level and the prior attainment of the students. Each dot in the figure represents a student and the line represents the average (or the trend) by mean GCSE score, obtained through non parametric smoothing. The grey ribbon around each line is the 95% confidence intervals of the average which provides an indication of the spread of data around

the average. Considering that there is no overlap between the confidence intervals for native and non-native speakers it is possible to conclude that, on average, native speakers perform significantly better than their native counterparts. However, the difference in performance varies according to prior attainment. For students with high mean GCSE scores (ie the higher achievers at GCSE) the gap tends to close rapidly, so that non-native speakers perform similarly to their native speaker counterparts.

From Figure 8 it is also possible to see that there is a wide variation around the average UMS achieved by students with similar values of mean GCSE. This is not necessarily surprising considering the many factors, in addition to prior attainment, that can influence attainment in specific subjects at A level. However, it highlights the importance of taking such factors into account when considering the performance of native and non-native speakers. To do this, a regression approach was used.

*French*



*German*



*Spanish*
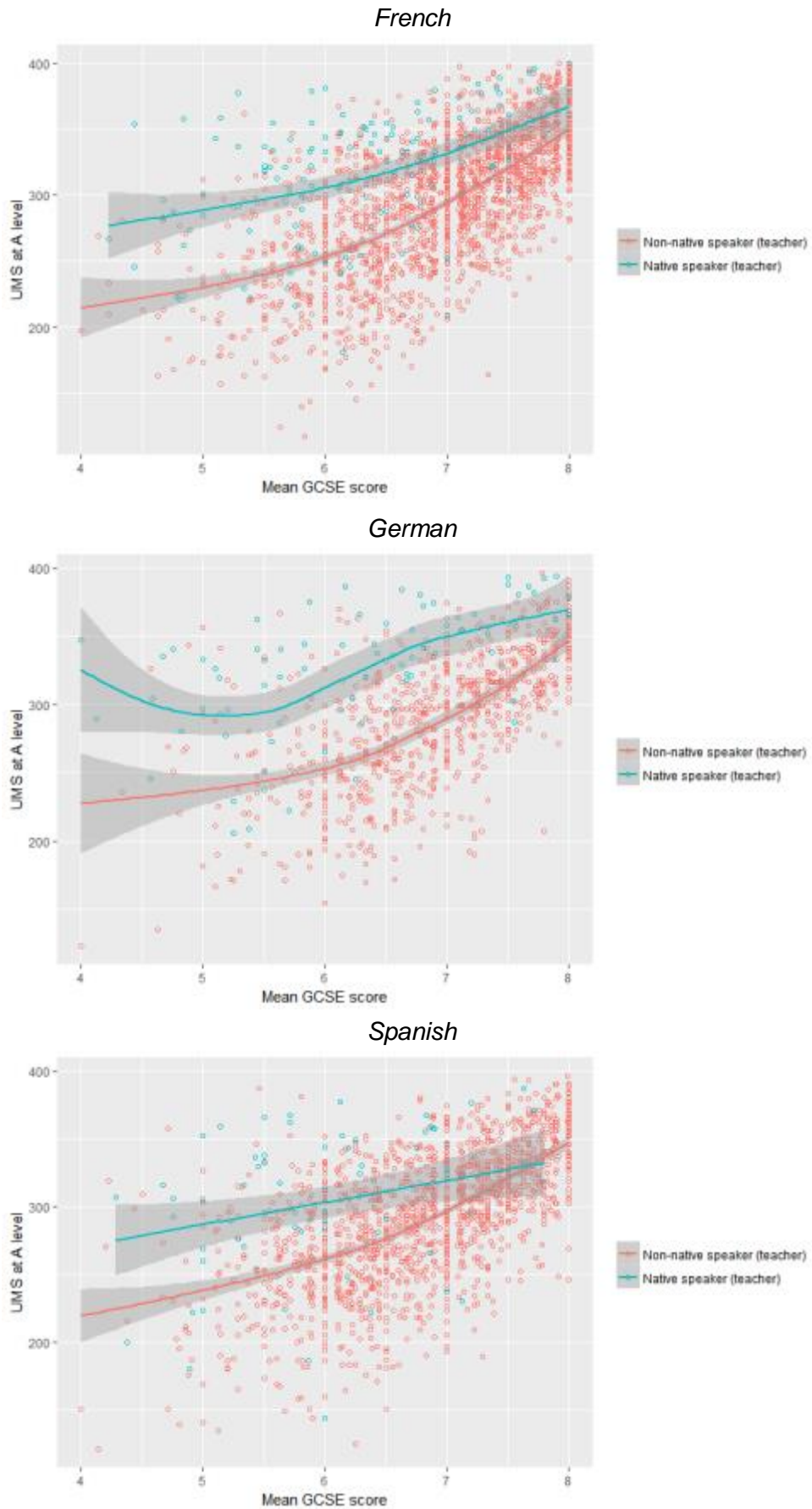


Figure 8. *UMS at A level and mean GCSE for native and non-native speakers*

For each of the 3 languages, Table 10 reports the estimates of 2 multilevel regression models that consider the effect of being a native speaker on the total UMS score when controlling for mean GCSE score, gender, and school type (see Appendix C for details of the modelling). The 2 models differ in that Model 2 includes an interaction term between mean GCSE score and whether students are a native speaker or not, allowing one to consider the effects of being a native speaker for students with different prior attainment.

Model 1 shows that after accounting for prior attainment, gender and school type, the association between whether students are native speakers or not and overall performance is statistically significant and quite large, ranging from nearly 44 UMS marks for Spanish, to nearly 56 UMS marks for German. This suggests that whether students are native speakers or not has a significant effect on performance, with native speakers outperforming their non-native speaker counterparts. The effect of gender and school type are also statistically significant, with girls and independent school students outperforming boys and state school students. However, these effects are much smaller.

Model 2 shows that for each language, there is also a significant interaction between mean GCSE score and whether students are native speakers. The interaction term is negative, indicating that as the mean GCSE score of a student increases, the difference in performance between native and non-native speakers decreases. In other words, the role of being a native speaker is larger for students with lower prior attainment. This is summarised in Table 11, which reports the effect of being a native speaker on UMS marks for students with a mean GCSE score of 5 (representing a typical grade C student at GCSE) and a mean GCSE of 7 (representing a typical grade A student at GCSE).

Table 10. *Multilevel regression for the UMS achieved at A level, by subject*

| Variable | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| | Est. | | S.E. | Est. | | S.E. |
| ***French*** | | | | | | |
| (Intercept) | 25.537 | *** | 7.914 | 12.947 | | 8.269 |
| Native speaker: yes | 46.714 | *** | 3.365 | 160.915 | *** | 23.306 |
| Mean GCSE score | 39.023 | *** | 1.152 | 40.799 | *** | 1.200 |
| School type: Independent | 10.132 | *** | 2.813 | 10.421 | *** | 2.799 |
| Gender: Female | -5.536 | *** | 2.089 | -5.210 | ** | 2.078 |
| Native speaker*Mean GCSE | - | | - | -17.867 | *** | 3.609 |
| ***German*** | | | | | | |
| (Intercept) | 56.800 | *** | 11.299 | 33.141 | *** | 12.691 |
| Native speaker: yes | 55.755 | *** | 4.208 | 152.055 | *** | 24.833 |
| Mean GCSE score | 33.992 | *** | 1.682 | 37.463 | *** | 1.884 |
| School type: Independent | 15.666 | *** | 4.290 | 14.610 | *** | 4.220 |
| Gender: Female | -6.953 | ** | 2.751 | -7.266 | *** | 2.730 |
| Native speaker*Mean GCSE | - | | - | -14.881 | *** | 3.787 |
| ***Spanish*** | | | | | | |
| (Intercept) | 55.826 | *** | 8.907 | 45.559 | *** | 9.140 |
| Native speaker: yes | 43.613 | *** | 4.778 | 174.785 | *** | 30.010 |
| Mean GCSE score | 34.981 | *** | 1.331 | 36.458 | *** | 1.363 |
| School type: Independent | 8.518 | *** | 3.283 | 8.885 | *** | 3.223 |
| Gender: Female | -6.397 | *** | 2.286 | -6.158 | *** | 2.273 |
| Native speaker*Mean GCSE | - | | - | -21.752 | *** | 4.909 |

*Note: These are the estimates of a random intercept regression model. Estimates are reported on the UMS scale. Mean GCSE is measured by converting grades into letters (A\*=8, A=7..). Reference category for school type is 'State schools', for gender is 'Male'. Stars indicate levels of significance: 0.01: \*\*\*; 0.05: \*\*; 0.1 \*.*

Table 11. *Differences in UMS between native speakers and non-native speakers, once gender, school type and mean GCSE are accounted for*

| | Model 1 | Model 2 | |
|---|---|---|---|
| | Average effect | Average GCSE grade C | Average GCSE grade A |
| French | 46.7 | 71.6 | 35.8 |
| German | 55.8 | 77.7 | 47.9 |
| Spanish | 43.6 | 66.0 | 22.5 |

*Note: Figures in the table are derived from Table 10. A grade C student is considered to have a mean GCSE of 5, a grade A student is considered to have a mean GCSE of 7.*

As achieving more UMS does not necessarily translate into a higher grade (this depends on how many UMS marks are achieved and whether this moves a student over the next grade boundary), further modelling was undertaken to explore the effect of being a native speaker on the probability of attaining certain grades. Table 12 summarises the average odds ratios associated with being a native speaker estimated through a series of multilevel logistic regression models fitted for the probability of attaining a grade A*, A or above, and C or above (the full results of the regression analysis are reported in Appendix E). The odds ratios allow us to interpret the effect of being a native speaker on the probability of attaining a certain grade, once other factors are controlled for (note that these analyses do not account for the interaction between prior attainment and the effects of being a native speaker). As an example, Table 12 shows that for A level French, native speakers were almost 5 times more likely to attain a grade A* than their non-native speaker counterparts when controlling for prior attainment, gender and school type. While the odds ratios are much smaller with respect to the probability of achieving a grade A (or above) and C (or above), they are still statistically significant.

For Spanish and German the odds ratios are much larger. For A level Spanish, native speakers are almost 10 times more likely to achieve a grade A (or above) or A* than non-native speakers, and 5 times more likely to achieve a grade C or above. For German the odds ratios are even greater: native speakers are 30 times more likely to achieve at least a grade C, 28 times more likely to achieve a grade A (or above), and 11 times more likely to achieve an A*. These results suggest that the effect of being a native speaker is greatest in German, but also relatively large for Spanish.

Table 12. *Odds ratio associated with being a native speaker for the probability of attaining a certain A level grade, by subject*

|  | A* | A or above | C or above |
|---|---|---|---|
| French | 4.94 | 1.40 | 1.35 |
| German | 11.12 | 28.11 | 31.99 |
| Spanish | 9.05 | 9.52 | 4.91 |

*Note: These are average odds ratios derived from multilevel logistic regression models and computed across students with different Mean GCSE. Standards errors not reported as all the estimates are statistically significant.*

The logistic regression models summarised in Table 12 show the average effect of being a native speaker, but do not consider the differential effects for students with different prior attainment. Figure 9 displays how the probability of achieving a grade A*, A (or above) or C (or above) changes according to mean GCSE score for native and non-native speakers, focusing on male students in state schools. This depicts a much more complex picture. For example, in A level Spanish, although native speakers perform better on average than non-native speakers, there is no difference

in the probability of attaining a grade C (or above) for students with high mean GCSE scores (6.5). At grade A*, non-native speakers with high prior attainment (greater than 7.5) perform better than native speakers, once gender and school type are accounted for. However, when interpreting these figures, it should be borne in mind that the proportion of students taking A level languages with such high levels of prior attainment is not very large, and that the number of native speakers is very small (Figure 6). Thus, these findings should be interpreted with caution.
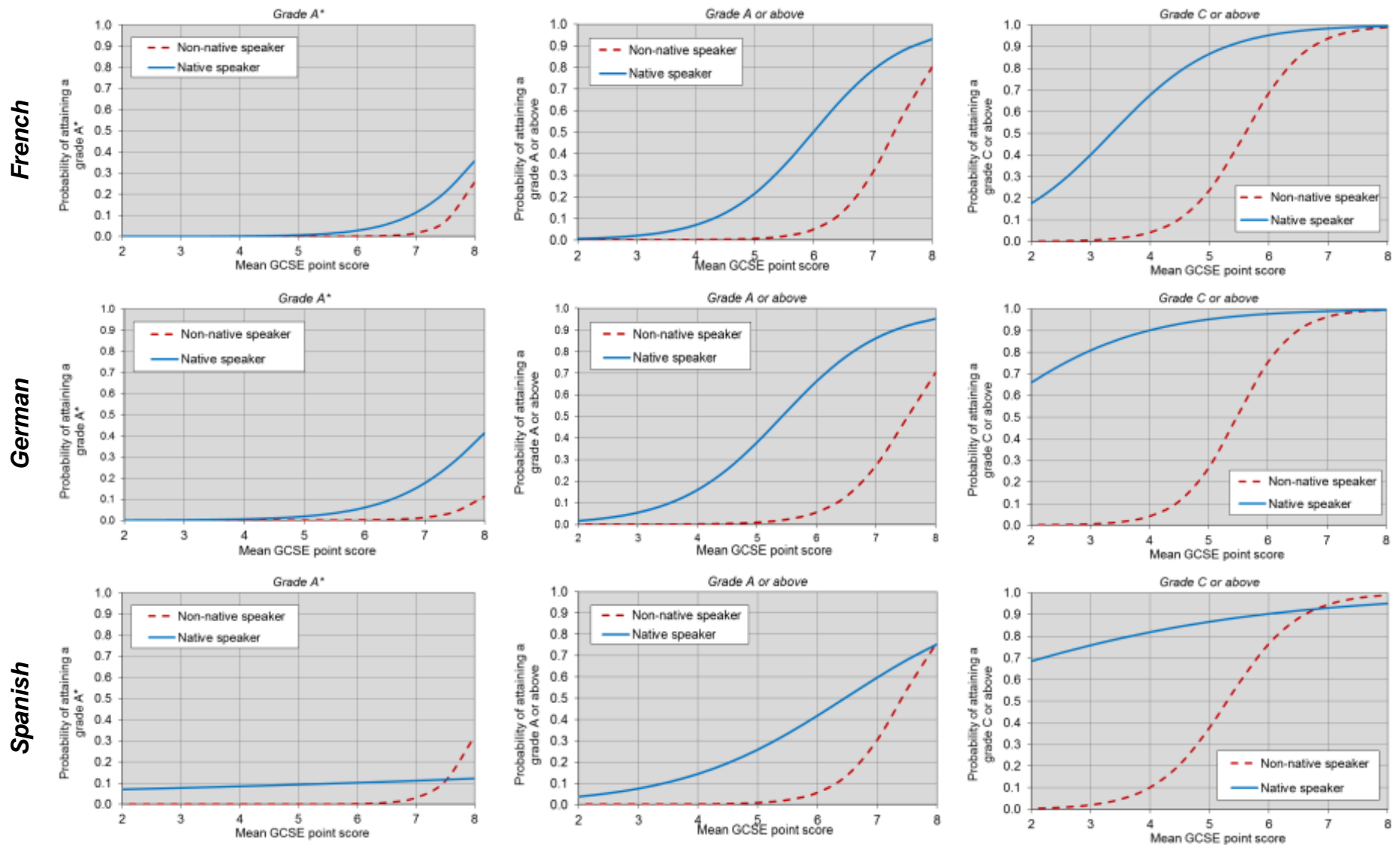
Figure 9. *The predicted probability of attaining a grade A\*/A/C (or above) by mean GCSE (for male students in state schools)*

## 5.5 Comparing the performance of native and non-native speakers at unit level

While the effects of being a native speaker are most important to consider at qualification level since the statistical predictions that drive the setting of grade boundaries are generated at qualification level, further analyses were conducted to consider the effect of being a native speaker at unit level. The units of the A level qualifications assess different skills, and it is possible that the effect of being a native speaker differs depending on the skills that are being assessed.

The A level MFL qualifications included in this research each contain four units, 2 at AS and 2 at A2, with one unit at each level assessing speaking, and the other unit assesses listening, reading, and writing[24]. Although the structure of the qualifications is similar across exam boards, the weighting and maximum unit marks are slightly different, so the unit level analyses only focused on the largest entry specification in each subject. In these specifications, the 2 speaking units have a weighting of 15% each and are assessed through discussion/conversation with the examiner. Listening, reading and writing amount to 70% of the overall assessment (35% at AS and 35% at A2) and are assessed with exam papers including different types of questions and essay-style pieces of writing, where students are assessed for both the content and the quality of language[25].

It could be argued that, since native speakers are likely to be exposed to the language mainly in their family environment, it is the speaking units where they might be more likely to outperform their non-native speaker counterparts. In order to check whether this hypothesis is true, it is possible to look at the performance of native and non-native speakers across the units that make up the A level MFL. However, due to the relatively small sample size, it is not possible to confidently rely on estimates of regression models. For this reason, only a descriptive investigation was performed to highlight differences between the performance of native and non-native speakers in terms of UMS marks across the different units and languages, similar to what has been shown in section 5.4. The results are displayed separately for French, German and Spanish, by Figures 10, 11 and 12, respectively.

It is interesting to note some patterns that are common to the 3 subjects. First, the positive relationship between mean GCSE score and UMS marks seems to be

---

[24] Note that the reformed A level qualifications, first assessed in summer 2018, generally contain three units: one assessing listening, reading and writing/translation, one assessing writing, and one assessing speaking.

[25] For an example, see the AQA A level French 2015 Unit 1 Writing paper (http://filestore.aqa.org.uk/subjects/AQA-FREN1-INS-JUN15.PDF) and mark scheme (http://filestore.aqa.org.uk/subjects/AQA-FREN1-W-MS-JUN15.PDF).

stronger in the reading, listening and writing units (1 and 3) than in the Speaking units (2 and 4), which may be due to the nature of the different assessments.

Second, native speaker students tend to outperform non-native speakers across the four units (with the exception of Spanish), though to a greater extent in the speaking units where many of them achieve the maximum UMS marks available[26]. This indicates that native speakers are particularly advantaged in the speaking elements of the assessment.

Third, the difference in the performance of native and non-native speakers is not constant across the range of prior attainment: as shown by the analysis at qualification level, non-native speakers with high mean GCSE score tend to close the gap and achieve the same UMS marks as their native speaker counterparts. This evidence suggests that, even in the speaking units, it is possible for non-native students to perform very well and achieve the maximum marks available.

---

[26] This suggests that there might be a 'ceiling effect', that is, native speakers' tend to achieve the maximum UMS mark available.
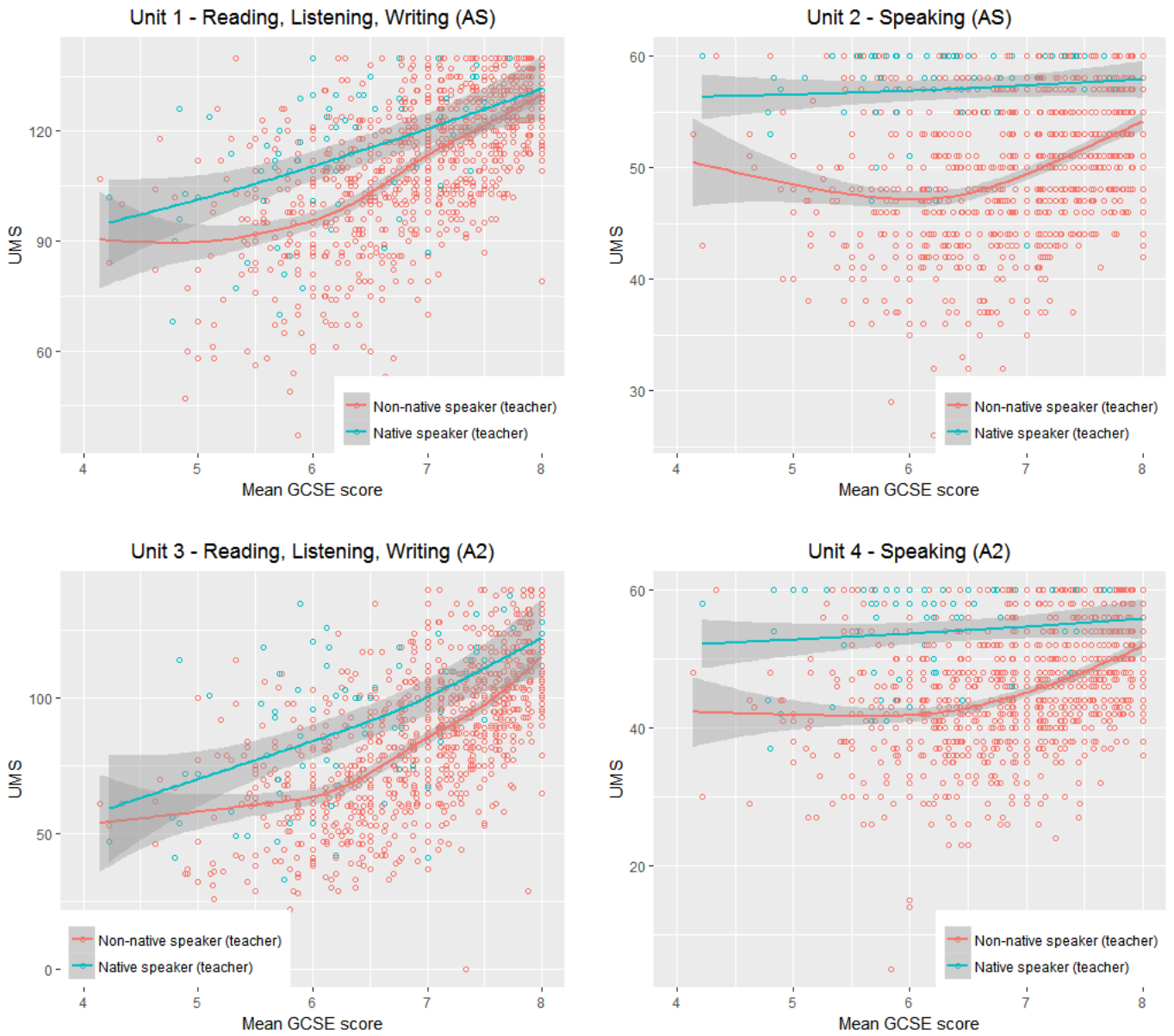
Figure 10. *The performance of native and non-native speakers at unit level - French*
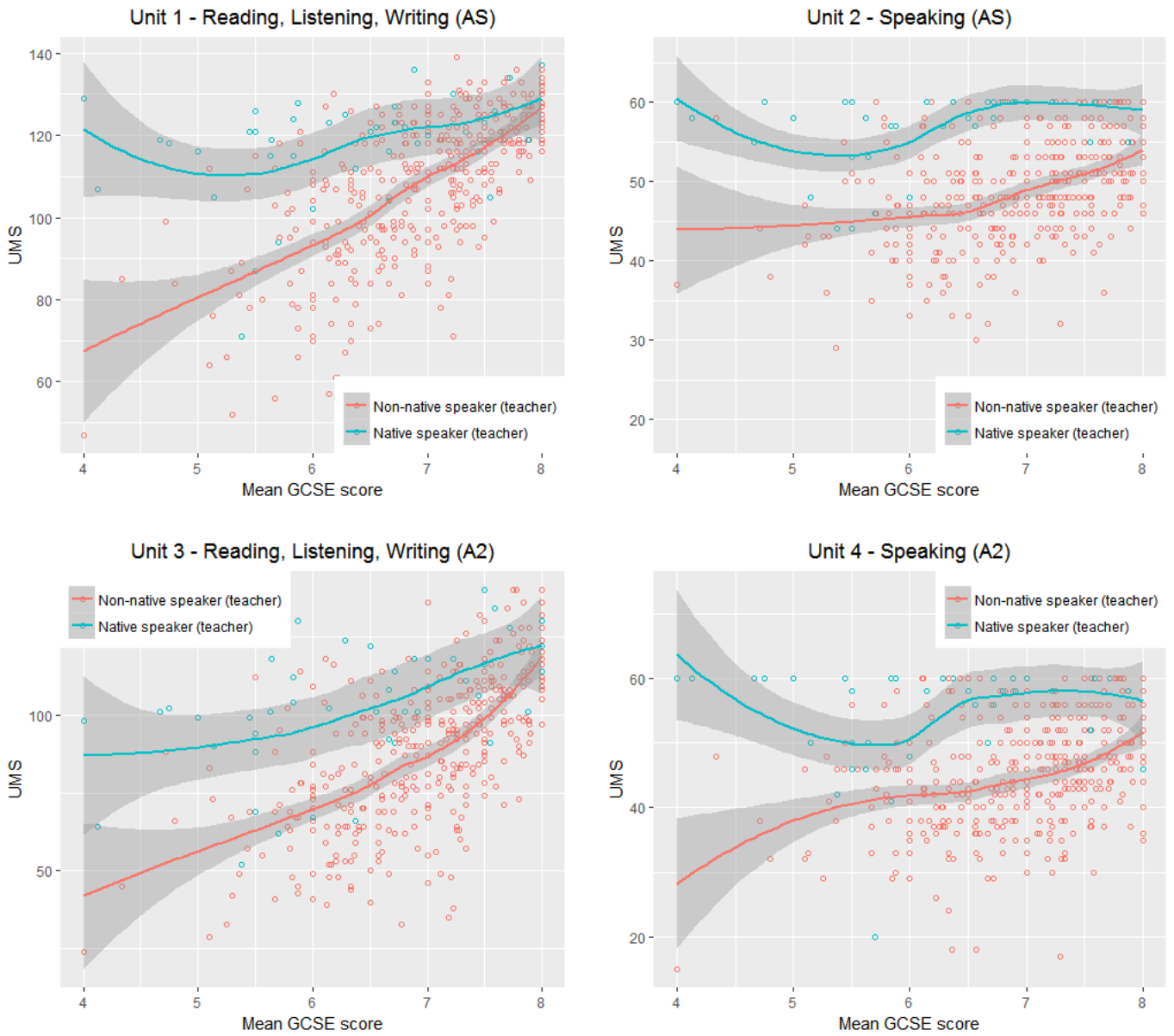
Figure 11. *The performance of native and non-native speakers at unit level - German*
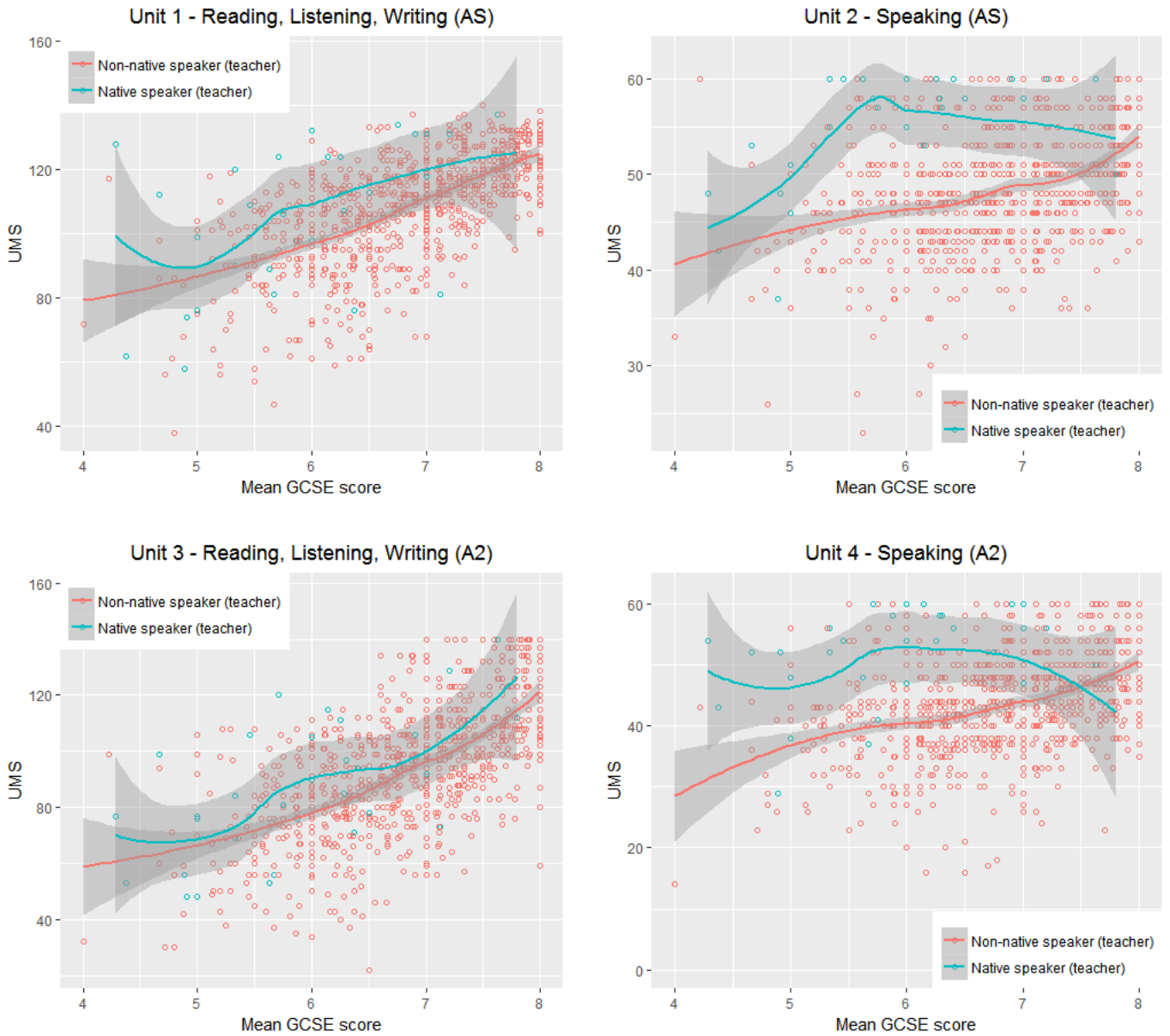
Figure 12. *The performance of native and non-native speakers at unit level - Spanish*

## 5.6    Implications for the maintenance of standards

The final analyses considered the implications of the presence of native speakers for maintaining standards, by considering how particular groups of students performed relative to their respective statistical prediction (see appendix A for a description of how the statistical predictions are generated). This allows an insight into whether, as a group, each set of students performed close to their prediction, above it, or below it. If a group performed above or below their prediction, it provides some indication that the relationship between prior attainment and outcomes is different for that group of students.

Predictions were generated for 3 groups of students in each A level MFL – students identified by their teacher as a native speaker; students identified by their teacher as a non-native speaker; and these 2 groups combined – using the same method that the exam boards use when setting grade boundaries[27]. These predictions were then compared to the outcomes for each group (ie the qualification grades achieved), with the key variable being the difference between the predictions and the outcomes. While the purpose of this research was to consider whether native speakers perform differently to non-native speakers, considering the native and non-native speakers combined together was intended to serve as a baseline measure against which the differences between the predictions and outcomes for the other 2 groups could be compared. For example, it would be possible for both native and non-native speakers to perform above or below their respective prediction, meaning that it is more informative to interpret the findings within the context of how the combined cohort perform.

Table 13 shows the predictions, outcomes, and the differences between the predictions and outcomes for each subject and group of students. A positive difference between the outcomes and predictions indicates that the students performed above prediction, while a negative difference indicates that the students performed below prediction. The analyses focus on grade A. This is a key grade boundary at A level that is set using statistical and judgemental evidence and is subject to reporting tolerances[28]. While grade E is also a key boundary set using statistical and judgemental evidence, it is not subject to reporting tolerances. Furthermore, since the majority of students are predicted to achieve at least a grade E, any differences between the predictions and outcomes at this grade are negligible. Nonetheless, the full results for all grades are provided in Appendix F.

Considering the combined native and non-native group first, Table 13 shows that the outcomes for this group were above (but very close to) prediction at grade A for all

---

[27] The analyses only include those students that are matched to their prior attainment, since this is one of the conditions for being included in the statistical predictions.
[28] Grade A* is also subject to reporting tolerances but is not set using examiner judgement.

subjects. This reflects the outcomes in the summer 2016 series: outcomes were close to (or above) the statistical predictions for each exam board individually (Ofqual, 2016c), and when combined across exam boards, the latter yielding outcomes above predictions for French (+0.72%), German (+1.28%), and Spanish (+1.81%). This suggests that the sample of students included in the research is fairly representative of the total candidature certificating in summer 2016, in terms of their performance relative to prediction.

The second group to consider are the native speakers. In doing so, it is important to consider that there are very few native speakers who are matched to their prior attainment (and can therefore be included in the predictions) in each subject. This is partly due to the response rate for this study (which was generally less than 30% of the total entry), but also due to the percentage of students that are native speakers and matched to their prior attainment being relatively low, eg in German around 10% of matched students were identified as native speakers. This means that the number of students included in the native speaker predictions are low, suggesting that the predictions (and therefore the comparisons with the outcomes) are likely to be less reliable. Indeed, the number of matched native speakers in each subject falls well below 500 students, the threshold below which exam boards generally prioritise alternative evidence when setting grade boundaries (Ofqual, 2016a). As such, the comparisons between the predictions and outcomes for the native speakers must be interpreted cautiously.

Table 13 shows that for each subject, the native speakers outperform their predictions by a considerable extent, particularly in German. This is not surprising given the previous findings in this paper that showed a statistically significant difference in performance between native and non-native speakers, once prior attainment, gender and school type have been controlled for. Although these analyses must be interpreted with some caution due to the low number of students, given the size of the differences and the consistency across subjects, it is unlikely that the effect is only attributable to the inherent unreliability associated with generating predictions for so few students. However, it is plausible that at least the size of the differences are influenced by this.

The final group of students to consider are the non-native speakers. These students make up the majority of respondents so there are sufficient students to generate reliable predictions in each subject. Table 13 shows that the outcomes for these students tend to be below prediction, particularly in German. Here, the outcomes are 4% below prediction. A similar (yet smaller) effect is seen in French and Spanish – in French, the outcomes for the non-native speakers were 1.21% below prediction, while in Spanish the outcomes for the non-native speakers were 0.77% below prediction. To provide some context, it is worth considering these differences alongside the reporting tolerances that are applied to A level outcomes each year (see Ofqual, 2016b). For A level French, German and Spanish these vary by exam board (due to differences in the number of matched students), and for some exam

boards, there were no tolerances at all in 2016 (see Ofqual, 2016c). Where there were tolerances, these were between 1%-3% for A level French, 2%-3% for A level German, and 1%-2% for A level Spanish (see Ofqual, 2016c, for the reporting tolerances by exam board for each subject in 2016). As such, for some subjects, the differences between the outcomes and predictions when excluding native speakers are lower than the tolerances for some exam boards, meaning that outcomes could deviate from predictions to this extent but not be out of tolerance.

Given that the outcomes for the native and non-native speakers combined were above prediction in each subject, it is also possible to compare the differences between predictions and outcomes for non-native speakers with the differences between predictions and outcomes for the 2 groups combined. This suggests that the inclusion of native speakers resulted in the difference between the predictions and the outcomes changing by 2.67% in French, 4.42% in German, and 1.20% in Spanish. In Spanish, the effects therefore seem relatively small, likely due to the relatively small number of 'matched' students that are native speakers (4.9%; see Table 8). The effects are greater in French and relatively large in German. This is not surprising given that the native speakers exceed their prediction in German to the greatest extent, and that German has the highest percentage of native speakers (10.9% of the students matched to their prior attainment were identified as native speakers by the teachers; Table 8).

Table 13. *A level French, Germans and Spanish outcomes vs. predictions – grade A*

| Subject | | All students | Native speakers | Non-native speakers |
|---|---|---|---|---|
| French | Prediction | 35.74 | 21.66 | 36.91 |
| | Outcome | 37.20 | 55.30 | 35.70 |
| | Difference | **1.46** | **33.64** | **-1.21** |
| | Number | 1,991 | 152 | 1,839 |
| German | Prediction | 31.38 | 26.93 | 31.90 |
| | Outcome | 31.80 | 65.20 | 27.90 |
| | Difference | **0.42** | **38.27** | **-4.00** |
| | Number | 849 | 89 | 760 |
| Spanish | Prediction | 31.17 | 17.59 | 31.87 |
| | Outcome | 31.60 | 42.10 | 31.10 |
| | Difference | **0.43** | **24.51** | **-0.77** |
| | Number | 1,556 | 76 | 1,480 |

Overall, these findings suggest that native and non-native speaker perform differently given their prior attainment (ie have a different rate of value added from mean GCSE score), with native speakers outperforming their prior attainment based predictions to a considerable extent. This provides further support for the previous findings in this paper that showed a significant difference in performance for native

and non-native speakers once prior attainment, gender and school type have been controlled for. However, given that there are relatively few native speakers in each subject, their inclusion in the statistical predictions has a relatively small effect in French and Spanish compared to German, where the effect is greater. Furthermore, given the challenges in identifying native speakers and the relatively low response rate in our research, the size of the effect can only be considered as an estimate.

While these analyses suggest an effect of including native speakers in the statistical predictions, it is important to note that the data collected here relate to only one examination series. Thus, it is not possible to determine whether the proportion of native speakers has changed over time and this is key when considering the impact of native speakers on the maintenance of standards. If the proportion of native speakers has remained stable over time then their inclusion in the statistical predictions should not be problematic (assuming that an appropriate standard was set initially), since they will be predicted to perform the same each series. However, if the number of native speakers has increased over time, the statistical predictions would not allow for this and the greater number of students with higher value added. Ascertaining the extent to which the percentage of native speakers has changed is not possible from this research though.

# 6 Discussion

## 6.1 Summary of findings

This research aimed to consider the presence of native speakers in A level MFL. The research focused on 5 MFL subjects – French, German, Spanish, Italian and Russian – and aimed to explore 3 research questions: i) what is the percentage of native speakers sitting each A level MFL; ii) how do native speakers perform on the A level assessments compared to non-native speakers, and iii) what are the potential implications for maintaining standards. The analyses in this report focus on A level French, German and Spanish due to the low response rate and the smaller entries for A level Italian and Russian.

There were four key findings. First, the research confirmed the challenges in identifying native speakers. The extent to which individuals are native or non-native speakers is likely to exist on a continuum, yet to satisfy the aims of this research, it was necessary to distinguish between native and non-native speakers dichotomously. Due to the anticipated challenges in doing this, 2 pieces of information were collected: students were asked to complete a questionnaire relating to their language experience and proficiency, and teachers were asked to identify whether students were native speakers or not in the language that they were studying at A level, with the intention of using both pieces of information to identify native speakers. However, a number of analyses showed that it was not possible to use the information provided by students to identify native speakers due to issues with the data collected, and the research relied on the responses from teachers instead.

Using this measure, the percentage of native speakers in each subject was estimated and the performance of native and non-native speakers on the assessment was compared. This showed that the percentage of native speakers varied by subject and was relatively small but not insignificant, particularly in A level German, where the percentage of native speakers was greatest. The percentage of native speakers was smaller when considering just those students that are matched to their prior attainment (ie those students that are included in the statistical predictions), than when considering all students. This suggests that any effect of native speakers on the maintenance of standards might be smaller than anticipated by stakeholders.

A number of regression analyses showed that native speakers outperform the non-native speakers on the overall A level assessment once their prior attainment, gender and school type have been controlled for. However, the size of the effect differed: the effect of being a native speaker on overall performance was greatest for the students with lower prior attainment, and greatest in A level German.

Native speakers were also found to outperform a prior attainment based prediction, particularly in A level German. This suggests that native speakers have a different value added from mean GCSE to non-native speakers. However, due to the relatively low number of native speakers that are included in the predictions, these findings must be interpreted with caution. Furthermore, the effect of removing native speakers from the statistical predictions was relatively small in French and Spanish, and in some cases, smaller than the tolerances that are applied to the outcomes for some exam boards.

## 6.2  Limitations

There are a number of limitations to consider when interpreting the findings of this research, the first (and perhaps most significant) relates to the identification of native speakers. As outlined above, identifying native speakers via the student questionnaire was not possible. This suggests that the student questionnaire that we used, an adapted version of the LEAP-Q, did not function as intended with our population of students. The original LEAP-Q (Marian, Blumenfeld & Kaushanskaya, 2007) was developed for identifying bilingualism and validated with individuals who have significant exposure to multiple languages. While many of the facets of bilingualism are similar to characteristics of native speakers, it is possible that the questionnaire did not function as effectively for some of our students who had limited exposure to their second language, ie those that only had exposure through learning the language at A level, despite us adapting the questionnaire to facilitate this.

Alternatively, students might have mis-interpreted some of the items on the questionnaire. Closer examination of the responses from individual students suggested that some responses were contradictory, suggesting that some students might have mis-interpreted the questions. Given this, one might suggest that the questionnaire should have been refined further prior to use in this research. However, the questionnaire was piloted with current year 13 students and circulated to a number of language associations for comment prior to distribution, so it is not obvious what further steps could have been taken to adapt the questionnaire. Indeed, it is not clear what other pieces of information could be collected from students to more reliably identify native speakers using a questionnaire approach, without relying on a precise definition of a native speaker (problematic in itself) that is operationalised in a set of questions.

The challenges of using the information reported by students meant that the research relied on the teachers' reports of whether students were native speakers or not. Although this was considered to be the most reliable method of identifying native speakers, this measure itself might have been problematic. It is possible that some teachers might have been less familiar with their students' familial background than others, or some teachers perceptions might have been influenced by the proficiency of their students (ie a highly proficient student might be assumed to have some native speaker characteristics). Although this represents a limitation to reliably

identifying native speakers, using the responses from teachers offers the advantage of allowing teachers' concerns regarding the presence of native speakers in the entry cohort for A level MFL, and their impact on the maintenance of standards, to be tested empirically.

The challenges in identifying native speakers has implications for any similar further research that might be conducted, since it is not clear how native speakers could be more effectively identified. Furthermore, there are implications for the other findings reported here. Using the responses from teachers, we were able to estimate the percentage of native speakers sitting each A level MFL and compare the performance of native and non-native speakers on the assessment and to prior attainment based predictions, information that has not previously been available. However, any figures resulting from these analyses can only be treated as estimates. They are based on only a small proportion of the overall A level cohort in each subject, and more importantly, they rely on the method used to identify native speakers (ie the teacher reports). If the sample was larger, or if an alternative method to identify native speakers was used, the figures (though not necessarily any conclusions) would almost certainly differ. Furthermore, the estimates of the percentage of native speakers provide no indication of whether this has changed over time, which has implications for the conclusions that can be drawn in relation to the maintenance of standards.

The final limitations are more methodological. First, although the research focused on the 3 larger entry subjects (French, German and Spanish), the response rate was still relatively low. Analyses suggested that our final samples were reasonably representative of the overall A level cohort in each subject in terms of student characteristics, but it might be possible that other factors that we cannot account for influenced the response rates. For example, some schools might have been more motivated to respond than others, eg those schools with a large number of native speakers. Furthermore, while efforts were made to reassure schools that no students would be advantaged or disadvantaged based on their participation, some schools or students might have been disinclined to respond due to such concerns. This might have biased our sample on characteristics that we are not able to compare. The small sample size also means that when the data is broken down by language, specification and/or unit, any statistics produced on native speakers is based on a very small number of students, which means that the findings might be influenced by noise in the data.

Second, while the analyses comparing the performance of native and non-native speakers used regression techniques that controlled for a student's prior attainment, gender, and school type, these are only a few of the possible variables that could influence performance that we have data on. It is likely that if more predictors of performance were included in the models, then the native speaker effect would be reduced. For this reason, caution must be taken when interpreting the results of the regression analyses presented in this report. The estimates retrieved can only

ascertain measures of associations, but cannot be taken as estimates of underlying causal relationships.

## 6.3    Conclusions

Despite the limitations, this research suggests that native and non-native speakers do perform differently on the A level assessment (at qualification level) and have different value added from mean GCSE, though the size of the native speaker effect differs across subjects. While this suggests that the presence of native speakers might have implications for the maintenance of standards, this very much depends on whether the proportion of native speakers has changed over time. If it has remained stable, there should be no effect of native speakers, assuming that the native speakers perform similarly each year and that an appropriate standard was set initially, since this will have been carried forward each subsequent year. If the proportion of native speakers has increased though, it is likely that the statistical predictions would have under-estimated their performance due to the changing nature of the cohort. Whether the proportion of native speakers has changed cannot be known from this research alone. However, there is reason to think that it might have given the increasing immigration of French, German and Spain nationals over recent years, the increasing proportion of EAL students taking A level MFL, and the general decline in entries in these subjects.

Although this evidence suggests that the inclusion of native speakers might have implications for maintaining standards, it must also be considered that the setting of grade boundaries uses a combination of statistical and judgemental evidence. As such, if the statistical predictions were significantly under-predicting outcomes then the senior examiners' judgement should be mis-aligned with the statistical evidence. However, this does not seem to have been the case, given that none of A level French, German and Spanish had outcomes that exceeded reporting tolerances in June 2015 or 2016. This might seem difficult to reconcile, but it is possible that any shifts in the standards each year have been very subtle, meaning that the limits of examiner judgement have prevented these changes from being observed (Ofqual, 2015), and that only a comparison between, say, 2016 and 2010, would highlight any changes in standards.

On balance, this research therefore suggests that there is likely to be a small, yet important effect, of native speakers in A level MFL. However, given the challenges of identifying native speakers, the research also suggests that routinely monitoring the presence of native speakers in A level MFL each year would not be possible, and attempts to do so would not be proportionate. It is therefore recommended that thought is given to whether an adjustment to the standards is appropriate.

# References

AQA (2013). GCE AS and A Level Specification: French/German/Spanish. Retrieved from http://filestore.aqa.org.uk/subjects/specifications/alevel/AQA-2650-2660-2695-W-SP-14.PDF

Alreck, P. L., and R. B. Settle. (1995). The Survey Research Handbook. 2nd ed. Boston, MA: Irwin/McGraw Hill.

Baird, J., Cresswell, M.J., & Newton, P. (2000). Would the *real* gold standard please step forward? *Research Papers in Education, 15*(2), 213-229.

Baker, P., & Eversley, J. (2000). *Multilingual capital: the languages of London's schoolchildren and their relevance to economic, social and educational policies*. London: Battlebridge, 2000.

Black, B. (2014). Chapter 4 A-Level Modern Foreign Languages : Assessment Instruments and Script Analysis, (July). Retrieved from http://www.jcq.org.uk/media-centre/news-releases/chapter-4---a-level-modern-foreign-languages-assessment-instruments-and-script-analysis

Board, K. and Tinsley, T. (2016). *Language Trends 2015/16. The state of language learning in primary and secondary schools in England*. Reading: Education Development Trust. Retrieved from https://www.britishcouncil.org/sites/default/files/language_trends_survey_2016_0.pdf

Coe, R. (2008). Comparability of GCSE examinations in different subjects: an application of the Rasch model. *Oxford Review of Education, 34,* 609-636.

Davies, A. (2003). *The native speaker: myth and reality* (2nd ed.). Bristol: Multilingual Matters.

DfE (2015). Modern Foreign Languages GCE AS and A level subject content. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/485569/GCE_A_AS_level_subject_content_modern_foreign_langs.pdf

Goldstein, H. (2011). *Multilevel Statistical Models (4th edition)*. Chichester: John Wiley & Sons.

HESA (2016). *Statistical First Release 2014/15: First year students by subject area, level, mode and sex*. Cheltenham: Higher Education Statistics Agency. Retrieved from: https://www.hesa.ac.uk/news/14-01-2016/sfr224-enrolments-and-qualifications

HESA (2017). *Statistical First Release 2015/16: First year students by subject area, level, mode and sex*. Cheltenham: Higher Education Statistics Agency.

Retrieved from: https://www.hesa.ac.uk/news/12-01-2017/sfr242-student-enrolments-and-qualifications

Ipsos Mori (2014). *Why is the take-up of modern foreign language A levels in decline?* A report for the Joint Council for Qualifications. London: Ipsos MORI. Retrieved from: http://www.jcq.org.uk/media-centre/news-releases/chapter-2---why-is-the-take-up-of-modern-foreign-language-a-levels-in-decline

JCQ (2016). GCE Trends 2016. Retrieved from: http://www.jcq.org.uk/examination-results/a-levels/2016/gce-trends-2016

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage, *Journal of Cognitive Psychology*, *25*, 605-621.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940–967.

Newton, P. (2011). *A level pass rates and the enduring myth of norm-referencing.* Research Matters: A Cambridge Assessment publication. Special Issue 2: Comparability, 20 – 23.

Ofqual (2014). Exploration of assessment quality issues in A level modern foreign languages – technical report. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/387729/2014-09-26-exploration-of-assessment-quality-issues-in-a-level-modern-foreign-languages-technical-report.pdf

Ofqual (2015). Setting GCSE, AS and A level standards in summer 2014 and 2015. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/451321/2015-08-05-summer-series-gcse-as-and-a-level-grade-standards.pdf

Ofqual (2016a). Data exchange document. Retrieved from: http://dera.ioe.ac.uk/26661/1/Summer_2016_Data_Exchange_Procedures.pdf

Ofqual (2016b). Summer 2016 GCSE, AS and A level awards: a summary of our monitoring. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/548237/Summer_2016_GCSE_AS_and_A_level_awards_a_summary_of_our_monitoring.pdf

Ofqual (2016c). Inter-board comparability in GCSEs and A levels: summer 2016 exams. Retrieved from: https://www.gov.uk/government/publications/ofqual-2016-summer-exam-series-report

Ofqual (2016d). Ofqual corporate plan 2016-2019. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/513152/Ofqual_corporate_plan_2016_to_2019.pdf

Ofqual (2016e). GCSE and A level enquiries about results – subject level analyses. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/526064/2015_EAR_subject-level_analyses.pdf

Oxford English Dictionary (2016). Oxford English Dictionary. Oxford: Oxford University Press.

Rampton, M. B. H. (1990). Displacing the 'native speaker': expertise, affiliation, and inheritance. *English Language Teaching Journal, 44*(2)*,* 97–102.

Shaughnessy, J. J., and E. B. Zechmeister. (1994). *Research Methods in Psychology. 3rd ed.* New York: McGraw Hill.

Smith, J., Alton, A. & Mitchell, T. (2012). *Comparative analysis of A level student work: Subject report.* A report commissioned by Ofqual. Coventry: Ofqual. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/377922/2012-11-02-comparative-analysis-of-a-level-student-work.pdf

Vidal Rodeiro, C.L. (2017). The study of foreign languages in England: uptake in secondary school and progression to higher education. *Language, Culture and Curriculum*, forthcoming.

# Appendix A – Generating statistical predictions

Prior attainment based predictions model the relationship between prior attainment and outcomes in a reference year, then use this relationship to predict the expected outcomes of students in the current year. At A level, prior attainment is measured as students' mean GCSE score. There are 2 main steps to generating predictions for each subject, as follows.

First, an 'outcome matrix' is generated for the reference year (see figure A1 for a fictitious outcome matrix). Students in the reference year that are matched to their prior attainment are divided into deciles based on their prior attainment at GCSE. A matrix is then created that shows how each decile went on to perform in each A level. The top decile (numbered 1 in Figure A1) would include the most able students (based on mean GCSE score), and the bottom decile would include the least able students. Once constructed, this matrix shows the probability of achieving a given grade for students in each decile.

|  |  | A level grade | | | | | | |
|---|---|---|---|---|---|---|---|---|
|  |  | A* | A | B | C | D | E | U |
| Mean GCSE decile | 1 | 70 | 75 | 80 | 85 | 90 | 95 | 100 |
|  | 2 | 65 | 70 | 75 | 80 | 85 | 90 | 100 |
|  | 3 | 60 | 65 | 70 | 75 | 80 | 85 | 100 |
|  | 4 | 55 | 60 | 65 | 70 | 75 | 80 | 100 |
|  | 5 | 50 | 55 | 60 | 65 | 70 | 75 | 100 |
|  | 6 | 45 | 50 | 55 | 60 | 65 | 70 | 100 |
|  | 7 | 40 | 45 | 50 | 55 | 60 | 65 | 100 |
|  | 8 | 35 | 40 | 45 | 50 | 55 | 60 | 100 |
|  | 9 | 30 | 35 | 40 | 45 | 50 | 55 | 100 |
|  | 10 | 25 | 30 | 35 | 40 | 45 | 50 | 100 |

Figure A1. *Example outcome matrix*

Using this outcome matrix, it is then possible to predict how students in the current year are expected to perform, given their own prior attainment. For example, using the outcome matrix above, 70% of students in decile 1 might be expected to get a grade A*, 75% might be expected to get an A* or A; 80% might be expected to get an A*, A or B; and so on. This is repeated for each decile and then aggregated together to form a prediction for the probability of achieving each grade. Given that the predictions reflect the prior attainment profile of the students entering for each board, one board might have a higher prediction than another if the prior attainment profile of the students is higher.

# Appendix B – Student questionnaire

**Language Experience and Proficiency Questionnaire**

**Part 1: Please answer all questions in this part**

| Student Code | |
|---|---|

1) Please list all the languages you know **in order of dominance** (please mark languages of equal dominance with a *).

| | |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 | |

2) Please list all the languages you know **in order of acquisition** (native language first)

| | |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 | |

3) Please list what percentage of the time you are **currently** and **on average** exposed to each language **outside of school and school work / homework** (your percentages should add up to 100%)

| List language here | | | | | |
|---|---|---|---|---|---|
| List % here | | | | | |

4) Please list what percentage of the time you would choose to read a text in each of your languages **outside of school and school work / homework** (your percentages should add up to 100%)

| List language here | | | | | |
|---|---|---|---|---|---|
| List % here | | | | | |

5) Please list what percentage of the time you would choose to speak in each of your languages if you were speaking with a person fluent in all of your languages **outside of school and school work / homework** (your percentages should add up to 100%)

| List language here | | | | | |
|---|---|---|---|---|---|
| List % here | | | | | |

6) Please state your normal country of residence (if different from the UK)………………
Please turn over

**Part 2: Please only answer this part if you are taking A level FRENCH**

**If not, please turn to part 3**

1) Please state the age (in years) you began acquiring FRENCH ………………..

2) Please list the number of years and months you have spent in each language environment **not including school trips**

|  | Years | Months |
|---|---|---|
| A country where FRENCH is spoken |  |  |
| A family where FRENCH is spoken |  |  |
| A school where FRENCH is spoken as a first language |  |  |

3) Please state your level of proficiency in speaking, writing, reading and understanding FRENCH

|  | Novice | Intermediate | Fluent/mother tongue |
|---|---|---|---|
| Speaking |  |  |  |
| Writing |  |  |  |
| Reading |  |  |  |
| Understanding |  |  |  |

4) Please rate how much the following factors contributed to you learning FRENCH **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

5) Please rate the extent to which you are currently exposed to FRENCH in the following contexts **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

Please turn over

**Part 3: Please only answer this part if you are taking A level GERMAN**

**If not, please turn to part 4**

1) Please state the age (in years) you began acquiring GERMAN ………………..

2) Please list the number of years and months you have spent in each language environment **not including school trips**

|  | Years | Months |
|---|---|---|
| A country where GERMAN is spoken |  |  |
| A family where GERMAN is spoken |  |  |
| A school where GERMAN is spoken as a first language |  |  |

3) Please state your level of proficiency in speaking, writing, reading and understanding GERMAN

|  | Novice | Intermediate | Fluent/mother tongue |
|---|---|---|---|
| Speaking |  |  |  |
| Writing |  |  |  |
| Reading |  |  |  |
| Understanding |  |  |  |

4) Please rate how much the following factors contributed to you learning GERMAN **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

5) Please rate the extent to which you are currently exposed to GERMAN in the following contexts **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

Please turn over

**Part 4: Please only answer this part if you are taking A level SPANISH**

**If not, please turn to part 5**

1) Please state the age (in years) you began acquiring SPANISH ………………

2) Please list the number of years and months you have spent in each language environment **not including school trips**

|  | Years | Months |
|---|---|---|
| A country where SPANISH is spoken |  |  |
| A family where SPANISH is spoken |  |  |
| A school where SPANISH is spoken as a first language |  |  |

3) Please state your level of proficiency in speaking, writing, reading and understanding SPANISH

|  | Novice | Intermediate | Fluent/mother tongue |
|---|---|---|---|
| Speaking |  |  |  |
| Writing |  |  |  |
| Reading |  |  |  |
| Understanding |  |  |  |

4) Please rate how much the following factors contributed to you learning SPANISH **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

5) Please rate the extent to which you are currently exposed to SPANISH in the following contexts **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

Please turn over

**Part 5: Please only answer this part if you are taking A level RUSSIAN**

**If not, please turn to part 6**

1) Please state the age (in years) you began acquiring RUSSIAN ………………..

2) Please list the number of years and months you have spent in each language environment **not including school trips**

|  | Years | Months |
|---|---|---|
| A country where RUSSIAN is spoken |  |  |
| A family where RUSSIAN is spoken |  |  |
| A school where RUSSIAN is spoken as a first language |  |  |

3) Please state your level of proficiency in speaking, writing, reading and understanding RUSSIAN

|  | Novice | Intermediate | Fluent/mother tongue |
|---|---|---|---|
| Speaking |  |  |  |
| Writing |  |  |  |
| Reading |  |  |  |
| Understanding |  |  |  |

4) Please rate how much the following factors contributed to you learning RUSSIAN **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

5) Please rate the extent to which you are currently exposed to RUSSIAN in the following contexts **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

Please turn over

**Part 6: Please only answer this part if you are taking A level ITALIAN**

**If not, this is the end of the questionnaire**

1) Please state the age (in years) you began acquiring ITALIAN ………………

2) Please list the number of years and months you have spent in each language environment **not including school trips**

|  | Years | Months |
|---|---|---|
| A country where ITALIAN is spoken |  |  |
| A family where ITALIAN is spoken |  |  |
| A school where ITALIAN is spoken as a first language |  |  |

3) Please state your level of proficiency in speaking, writing, reading and understanding ITALIAN

|  | Novice | Intermediate | Fluent/mother tongue |
|---|---|---|---|
| Speaking |  |  |  |
| Writing |  |  |  |
| Reading |  |  |  |
| Understanding |  |  |  |

4) Please rate how much the following factors contributed to you learning ITALIAN **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

5) Please rate the extent to which you are currently exposed to ITALIAN in the following contexts **outside of school and school work / homework**

|  | Not at all | Minimally | Moderately | A lot |
|---|---|---|---|---|
| Interacting with friends |  |  |  |  |
| Interacting with family |  |  |  |  |
| Reading |  |  |  |  |
| Language CD/self-instruction |  |  |  |  |
| Watching TV |  |  |  |  |
| Listening to the radio/music |  |  |  |  |

End of questionnaire

# Appendix C – Analysis techniques

## Factor analysis

The first strand of analysis, aimed at quantifying the number of native speakers sitting A level MFL, involved performing factor analysis on data provided by students. Factor analysis is a statistical method that can be used to derive one or more latent unobserved variables called factors. The underlying idea of this method is that a small number of factors, though unobservable, are reflected into a larger set of indicators that can be observed. In other words, through the analysis of the variability of and correlation among the observable indicators, factor analysis can lead to the identification of the latent factor(s) actually generating the observed data.

In the case at hand, students were asked to report 2 sets of items on (i) the factors contributing to learning the language (part 2, question 4 of the questionnaire) and (ii) the kind of exposure they are currently experiencing to the language outside school and school related activities (part 2, question 5 of the questionnaire). This data was exploited through factor analysis in order to identify whether, and if so how, a method of identifying native speakers could be derived.

There are 2 types of factor analysis that can be used, exploratory and confirmatory. The latter is used when a theoretical model on the link between factors and indicators exists and has to be verified; the former is employed when no beliefs on the process generating the data exist. An exploratory approach was used in this study. This allowed us to identify whether the indicators thought to reflect native speaker characteristics were underpinned by the language exposure and the language background of students. If no factor was identified by this approach, this would have to be taken as evidence that the observed indicators are not linked to the latent variable we are seeking.

## Multilevel regression modelling

The second strand of analysis was aimed at comparing the performance of native and non-native speakers in A level MFL. In doing so, it is important to consider that native speakers can be different from non-native speakers with respect to not only their performance in A level MFL, but also other characteristics influencing their performance at A level. Therefore, a rough comparison of the results attained by native and non-native speakers could be biased.

To avoid this issue, regression modelling was used in order to explore the relationship between performance at A level and being a native speaker, once other factors were controlled for. In particular, the other factors controlled for were: prior attainment at GCSE (measured by students' mean GCSE score), gender and school type (independent or state-maintained). In this way, estimates of the regression coefficients can be interpreted as the relationship between each factor and performance, once the other factors are fixed.

It should be noted that a multilevel modelling approach was taken in order to account for the hierarchical structure of the data, as students are clustered within schools. If we failed to recognise this, then the standard errors of the regression coefficients would be underestimated, leading to an overestimate of the statistical significance of the parameter we are interested in (Goldstein, 2011).

Two measures of A level performance were used to compare native and non-native speakers: UMS mark and grade attained. In the former case the UMS mark of student $i$ in school $j$ was used as dependent variable in this specification of the regression model:

$$UMS_{ij} = \beta_0 + \beta_1 Native_{ij} + \beta_2 GCSE\ mean_{ij} + \beta_3 Gender_{ij} + \beta_4 School\ type + u_j + \varepsilon_{ij};$$

where: $\beta$. are the regression coefficients; $u_j$ is a random variable at school level, following a normal distribution with mean zero; $\varepsilon_{ij}$ is an error term. It should be noted that in this specification it is $\beta_1$ that is the main parameter of interest as it yields the relationship between performance and being native speaker.

In the latter case (when the role of being a native speaker on the grade attained at A level was investigated), 3 dependant variables were considered: whether a grade A*, a grade A (or above) and a grade C (or above) was achieved. In this case a logit regression model[29] was necessary to account for the dichotomous nature of the dependant variable (0 = grade not achieved; 1 = grade achieved). These models take the following form:

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right)_{ij} = \beta_0^* + \beta_1^* Native_{ij} + \beta_2^* GCSE\ mean_{ij} + \beta_3^* Gender_{ij} + \beta_4^* School\ type + u_j;$$

where: $p_{ij}$ is the probability of student $i$ in school $j$ achieving an A*, A (or above) or a C (or above). Also in this case, $\beta_1^*$ is the main parameter of interest: it yields the association between performance and being native speaker, once other factors are controlled for.

In addition to the regression coefficient, predicted probabilities and selected odds ratios are reported and discussed. In the context of this analysis, an odds ratio represents the increase in the odds of achieving a certain grade for a native speaker. Although the actual magnitude of the odds ratios is difficult to interpret, the relative magnitude of the odds ratios can be easily interpreted. An odds ratio greater than 1 indicates an increase in the likelihood of achieving a certain grade, with a greater odds ratio indicating a greater likelihood. Conversely, an odds ratio less than 1

---

[29] Alternatively, a multinomial regression could have been used to model simultaneously the probability of achieving different grades. However, using a set of logit regression models for the probability of attaining a certain grade allowed us to gain easiness of interpretation, without any loss of generality.

indicates a decrease in the likelihood of attaining a certain grade, with a smaller odds ratio indicating a smaller likelihood. An odds ratio equal to 1 indicates an equal likelihood of attaining a certain grade.

For both UMS mark and grade achieved, an alternative regression specification including the interaction term between mean GCSE score and whether a student was a native speaker was also fitted. This allows to check whether, and if so how, the difference in the performance at A level MFL between native speakers and non-native speakers varies according to the prior attainment of the students.

# Appendix D – Representativeness of the sample

Most of the analysis undertaken in this report focuses on a small proportion of the entries to A level MFL in summer 2016 (around one fourth). This is mainly due to the low response rate of schools, students and teachers taking part in our survey. In particular, the investigation of the native speaker effect on achievement and maintaining standards in A level MFL relies on students that the teachers reported as native speakers or not. In order to draw robust conclusions from this analysis, it is crucial that the sample for those we hold this information is representative of the entire population of interest. In practice, this means that there is a need to check whether the sample of students included in the analysis can be considered as representative of the entire candidature of A level French, Spanish and German.

A first piece of descriptive evidence on the representativeness of the sample is provided by the geographical location of the schools. Although schools in England, Wales and Northern Ireland were included in our study, it was only possible to retrieve the postcode of schools in England. Figure D1 shows the geographical location of all schools entering students to A level French, German and Spanish in summer 2016 (panel a) compared with those included in our sample (panel b), with the size of each point proportional to the number of students in that school who were entered to A level in the 3 languages or responded to the questionnaire. The comparison of the 2 maps highlights some differences between the whole MFL entry size and our sample, but shows that students in our sample are scattered across the country. It is also apparent that the provision of A level MFL is particularly high in London, and this seems to be reflected in our sample.

A second piece of evidence on the representativeness of the sample can be provided by investigating the probability of being included in our sample based on a number of predictors. This amounts to the estimation of logistic regression models, one for each language, where the dependant variable is whether or not students are included in our sample and the independent variables are those provided by AOs and available for the whole cohort of students: gender, mean GCSE score, grade achieved at A level, school type, and the awarding organisation providing the specification. The estimates of the logistic regression models, one for each language, are reported in Table D1.
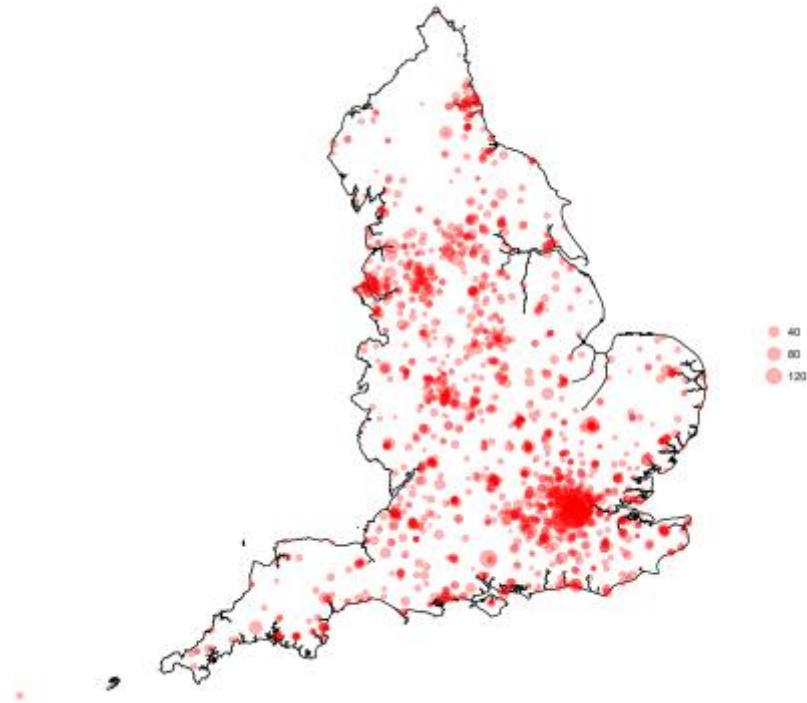
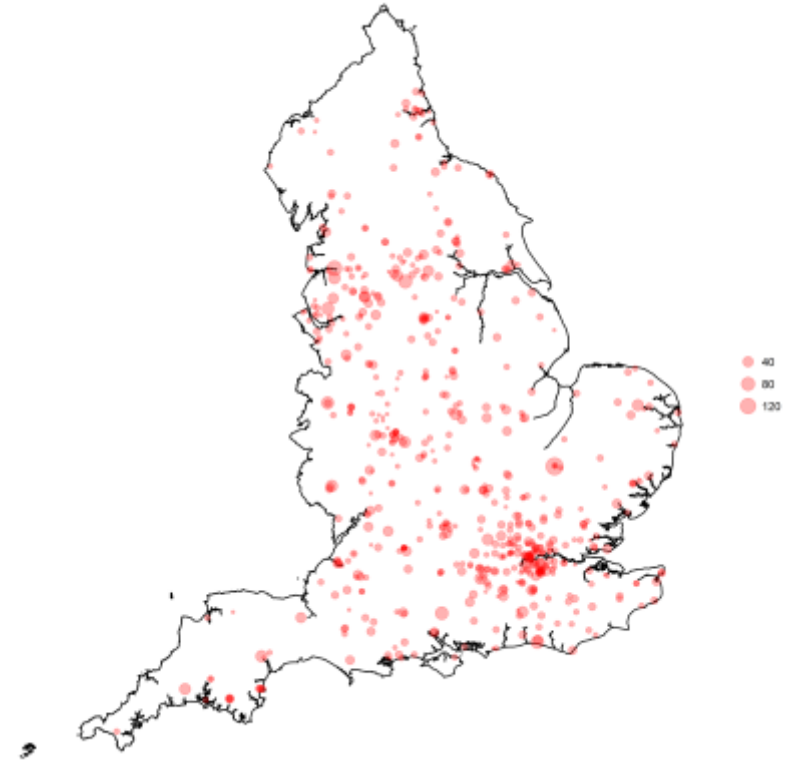*a. All MFL entries*    *b. Schools in our sample*



Figure D1.    *The geographical location of the schools providing A level MFL and of those in our sample*

Table D1.	*Logistic regression for the probability of being included in the analysis (ie students that teachers reported as a native or non-native speaker)*

| Variable | | French | | Spanish | | German | |
|---|---|---|---|---|---|---|---|
| | | Est. | S.E. | Est. | S.E. | Est. | S.E. |
| | (Intercept) | -1.7420*** | 0.418 | -0.9435 * | 0.4103 | -0.7922 | 0.6753 |
| Mean GCSE | score | 0.0004 | 0.042 | -0.0167 | 0.0453 | -0.0231 | 0.0594 |
| A level grade: | A* | -0.1629 | 0.107 | 0.0510 | 0.1244 | -0.2762 | 0.1869 |
| | A or above | 0.0447 | 0.072 | -0.0478 | 0.0838 | -0.1178 | 0.1121 |
| | B or above | 0.0572 | 0.080 | 0.0662 | 0.0853 | 0.0854 | 0.1203 |
| | C or above | 0.1711 | 0.104 | 0.0290 | 0.1119 | 0.0870 | 0.1584 |
| | D or above | 0.2278 | 0.159 | 0.1934 | 0.1871 | -0.0663 | 0.2501 |
| | E or above | 0.2346 | 0.376 | -0.4259 | 0.3659 | -0.0309 | 0.6271 |
| School: | Independent | 0.0486 | 0.068 | 0.0154 | 0.0754 | -0.0651 | 0.1118 |
| Gender: | Female | 0.2583*** | 0.061 | 0.2637 *** | 0.0667 | 0.2461 *** | 0.0892 |
| AO: | CCEA | -1.8497*** | 0.203 | -2.5393 *** | 0.3090 | -15.7012 | 265.4453 |
| | OCR | -0.1402 | 0.108 | 0.7786 *** | 0.1234 | -0.1222 | 0.1704 |
| | Pearson | -0.4229*** | 0.082 | -0.1728 * | 0.0885 | -0.5067 *** | 0.1212 |
| | WJEC | 0.3279*** | 0.068 | 0.3342 *** | 0.0804 | 0.4736 *** | 0.1036 |

*Note: Regression estimates based on 'matched' students. Reference category for school type is 'state schools', for gender is 'male', and for exam board is 'AQA'. Stars indicate levels of significance: 0.01: ***; 0.05: **; 0.1 *.*

Overall, findings on the representativeness of the sample are reassuring. Students of different levels of ability, measured in terms of both prior attainment (mean GCSE score) and grade achieved at A level, are equally represented in the sample considered for the analysis. This is an important result as it points towards the conclusion that, although a degree of caution must be applied in drawing inferences from such a small sample, there is no evidence to suggest that our sample is not representative of the whole candidature. The type of school attended was also found not to be over/under-represented in our sample.

Across the 3 subjects, only gender was highlighted as a significant predictor, with female students significantly more likely to take part in the survey. It is quite difficult to conceive how this could undermine the results related to the native speaker effect on performance and the maintenance of standards. The same consideration applies to the results associated to the provider of the specification they took. Although some AOs are over or under represented among our sample, there is no obvious indication of how this might affect the interpretation of the results.

# Appendix E – Additional regression estimates

This appendix shows the full estimates of the multilevel logistic regression analysis for the probability of attaining a grade A*, A or above, and C or above at A level.

Table E1 reports the estimate of the regression coefficients from which the odds ratios reported in the text (Table 12) have been computed. Consistent with the other regression models presented in this report, the independent variables included here are: being native speaker, prior attainment (measured through mean GCSE score), gender and school type.

Regression models in Table E1 do not account for the possible differential effect of being a native speaker given a student's prior attainment. This is accounted for in the regression models presented in Table E2 through the inclusion of an interaction term between mean GCSE score and being a native speaker. The predicted probabilities computed from these models are displayed in the main text by Figure 9.

Table E1. *Multilevel logistic regression for the probability of attaining at least a grade A\*, A and C at A level, by subject – Model 1*

| Variable | Grade A\* or above | | | Grade A or above | | | Grade C or above | | |
|---|---|---|---|---|---|---|---|---|---|
| | Est. | | S.E. | Est. | | S.E. | Est. | | S.E. |
| *French* | | | | | | | | | |
| (Intercept) | -22.576 | \*\*\* | 1.901 | -15.136 | \*\*\* | 0.855 | -10.603 | \*\*\* | 0.786 |
| Native speaker: yes | 1.599 | \*\*\* | 0.357 | 2.642 | \*\*\* | 0.268 | 2.606 | \*\*\* | 0.425 |
| Mean GCSE score | 2.677 | \*\*\* | 0.248 | 2.055 | \*\*\* | 0.119 | 1.904 | \*\*\* | 0.126 |
| School type: Independent | 0.371 | | 0.219 | 0.397 | \*\* | 0.179 | 0.639 | \*\*\* | 0.246 |
| Gender: Female | 0.189 | | 0.246 | -0.330 | \*\* | 0.146 | -0.286 | | 0.186 |
| Native speaker\*Mean GCSE | - | | - | - | | - | - | | - |
| *German* | | | | | | | | | |
| (Intercept) | -15.948 | \*\*\* | 2.310 | -12.750 | \*\*\* | 1.132 | -10.657 | \*\*\* | 1.472 |
| Native speaker: yes | 2.408 | \*\*\* | 0.395 | 3.336 | \*\*\* | 0.384 | 3.466 | \*\*\* | 0.724 |
| Mean GCSE score | 1.699 | \*\*\* | 0.309 | 1.684 | \*\*\* | 0.160 | 1.972 | \*\*\* | 0.243 |
| School type: Independent | 0.563 | | 0.348 | 0.691 | \*\*\* | 0.240 | 1.339 | \*\*\* | 0.585 |
| Gender: Female | 0.475 | | 0.399 | -0.429 | \*\* | 0.202 | -0.491 | \* | 0.294 |
| Native speaker\*Mean GCSE | | - | - | - | | - | | - | - |
| *Spanish* | | | | | | | | | |
| (Intercept) | -18.225 | \*\*\* | 1.716 | -14.040 | \*\*\* | 0.906 | -8.276 | \*\*\* | 0.833 |
| Native speaker: yes | 2.203 | \*\*\* | 0.511 | 2.254 | \*\*\* | 0.347 | 1.592 | \*\*\* | 0.450 |
| Mean GCSE score | 2.149 | \*\*\* | 0.230 | 1.889 | \*\*\* | 0.128 | 1.583 | \*\*\* | 0.136 |
| School type: Independent | 0.322 | | 0.264 | 0.510 | \*\*\* | 0.189 | 0.690 | \*\* | 0.296 |
| Gender: Female | -0.089 | | 0.267 | -0.242 | | 0.163 | -0.213 | | 0.195 |
| Native speaker\*Mean GCSE | - | | - | - | | - | - | | - |

*Note: These are the estimates of a random intercept model for the probability of attaining a certain grade. Estimates are reported on the logit scale (therefore not on the UMS scale). Mean GCSE is measured in score points retrieved converting grades into letters (A\*=8, A=7, …). Reference category for school type is 'State schools', for gender is 'Male'. Stars indicate levels of significance: 0.01: \*\*\*; 0.05: \*\*; 0.1: \*.*

Table E2. *Multilevel logistic regression for the probability of attaining at least a grade A\*, A and C at A level, by subject – Model 2, with interaction term*

| Variable | Grade A* or above | | | Grade A or above | | | Grade C or above | | |
|---|---|---|---|---|---|---|---|---|---|
| | Est. | | S.E. | Est. | | S.E. | Est. | | S.E. |
| *French* | | | | | | | | | |
| (Intercept) | -25.106 | *** | 2.237 | -16.000 | *** | 0.929 | -10.908 | *** | 0.814 |
| Native speaker: yes | 12.725 | *** | 3.740 | 8.186 | *** | 1.865 | 7.098 | *** | 2.570 |
| Mean GCSE score | 3.004 | *** | 0.291 | 2.174 | *** | 0.129 | 1.950 | *** | 0.130 |
| School type: Independent | 0.403 | * | 0.221 | 0.396 | ** | 0.180 | 0.642 | *** | 0.246 |
| Gender: Female | 0.207 | | 0.247 | -0.318 | ** | 0.147 | -0.280 | *** | 0.185 |
| Native speaker*Mean GCSE | -1.530 | *** | 0.515 | -0.869 | *** | 0.287 | -0.812 | * | 0.446 |
| *German* | | | | | | | | | |
| (Intercept) | -19.817 | *** | 3.540 | -13.889 | *** | 1.300 | -11.714 | *** | 1.595 |
| Native speaker: yes | 9.884 | ** | 4.467 | 7.584 | *** | 2.092 | 10.821 | *** | 3.287 |
| Mean GCSE score | 2.220 | *** | 0.470 | 1.845 | *** | 0.183 | 2.139 | *** | 0.263 |
| School type: Independent | 0.458 | | 0.348 | 0.650 | *** | 0.239 | 1.303 | ** | 0.587 |
| Gender: Female | 0.444 | | 0.394 | -0.437 | ** | 0.202 | -0.516 | * | 0.296 |
| Native speaker*Mean GCSE | -1.022 | * | 0.605 | -0.681 | ** | 0.325 | -1.363 | ** | 0.558 |
| *Spanish* | | | | | | | | | |
| (Intercept) | -22.265 | *** | 2.114 | -14.840 | *** | 0.160 | -8.891 | *** | 0.875 |
| Native speaker: yes | 19.499 | *** | 3.145 | 10.174 | *** | 1.895 | 8.936 | *** | 2.518 |
| Mean GCSE score | 2.690 | *** | 0.279 | 2.002 | *** | 0.002 | 1.679 | *** | 0.143 |
| School type: Independent | 0.312 | | 0.260 | 0.499 | *** | 0.186 | 0.707 | ** | 0.295 |
| Gender: Female | -0.032 | | 0.268 | -0.245 | | 0.161 | -0.207 | | 0.195 |
| Native speaker*Mean GCSE | -2.591 | *** | 0.476 | -1.279 | *** | 0.308 | -1.311 | *** | 0.426 |

*Note: These are the estimates of a random intercept model for the probability of attaining a certain grade. Estimates are reported on the logit scale (therefore not on the UMS scale). Mean GCSE is measured in score points retrieved converting grades into letters (A\*=8, A=7, …). Reference category for school type is 'State schools', for gender is 'Male'. Stars indicate levels of significance: 0.01: \*\*\*; 0.05: \*\*; 0.1: \*.*

# Appendix F – Outcomes compared to predictions

Table F1. *A level French outcomes and predictions*

| Natives speakers (n = 152) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 3.35 | 21.66 | 47.51 | 72.22 | 89.73 | 97.47 | 100.00 |
| Outcomes | 11.20 | 55.30 | 78.30 | 94.10 | 99.30 | 100.00 | 100.00 |
| Difference (outcome – prediction) | 7.85 | 33.64 | 30.79 | 21.88 | 9.57 | 2.53 | 0.00 |
| | | | | | | | |
| Non-native speakers (n = 1,839) | A* | A | B | C | D | E | U |
| Prediction | 6.83 | 36.91 | 66.22 | 85.23 | 95.51 | 99.16 | 100.00 |
| Outcomes | 7.80 | 35.70 | 65.90 | 85.60 | 95.60 | 99.50 | 100.00 |
| Difference (outcome – prediction) | 0.97 | -1.21 | -0.32 | 0.37 | 0.09 | 0.34 | 0.00 |
| | | | | | | | |
| Native + non-native speakers (n = 1,991) | A* | A | B | C | D | E | U |
| Prediction | 6.57 | 35.74 | 64.79 | 84.24 | 95.07 | 99.03 | 100.00 |
| Outcomes | 8.00 | 37.20 | 66.80 | 86.30 | 95.90 | 99.50 | 100.00 |
| Difference (outcome – prediction) | 1.43 | 1.46 | 2.01 | 2.06 | 0.83 | 0.47 | 0.00 |

Table F2. *A level German outcomes and predictions*

| Natives speakers (n = 89) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 4.48 | 26.93 | 50.90 | 72.86 | 89.11 | 97.17 | 100.00 |
| Outcomes | 18.00 | 65.20 | 85.40 | 94.40 | 100.00 | 100.00 | 100.00 |
| Difference (outcome – prediction) | 13.52 | 38.27 | 34.50 | 21.54 | 10.89 | 2.83 | 0.00 |

| Non-native speakers (n = 760) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 5.35 | 31.90 | 60.12 | 82.49 | 94.65 | 99.20 | 100.00 |
| Outcomes | 4.20 | 27.90 | 61.30 | 84.20 | 95.10 | 99.50 | 100.00 |
| Difference (outcome – prediction) | -1.15 | -4.00 | 1.18 | 1.71 | 0.45 | 0.30 | 0.00 |

| Native + non-native speakers (n = 849) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 5.25 | 31.38 | 59.15 | 81.48 | 94.07 | 98.99 | 100.00 |
| Outcomes | 5.70 | 31.80 | 63.80 | 85.30 | 95.60 | 99.50 | 100.00 |
| Difference (outcome – prediction) | 0.45 | 0.42 | 4.65 | 3.82 | 1.53 | 0.51 | 0.00 |

Table F3. *A level Spanish outcomes and predictions*

| Natives speakers (n = 76) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 2.63 | 17.59 | 41.83 | 68.72 | 87.55 | 96.42 | 100.00 |
| Outcomes | 11.80 | 42.10 | 76.30 | 86.80 | 96.10 | 98.70 | 100.00 |
| Difference (outcome – prediction) | 9.17 | 24.51 | 34.47 | 18.08 | 8.55 | 2.28 | 0.00 |

| Non-native speakers (n = 1,480) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 6.54 | 31.87 | 61.01 | 82.62 | 94.41 | 98.94 | 100.00 |
| Outcomes | 7.70 | 31.10 | 62.80 | 84.90 | 95.90 | 99.10 | 100.00 |
| Difference (outcome – prediction) | 1.16 | -0.77 | 1.79 | 2.28 | 1.49 | 0.16 | 0.00 |

| Native + non-native speakers (n = 1,556) | A* | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| Prediction | 6.35 | 31.17 | 60.07 | 81.94 | 94.08 | 98.82 | 100.00 |
| Outcomes | 7.90 | 31.60 | 63.40 | 85.00 | 95.90 | 99.10 | 100.00 |
| Difference (outcome – prediction) | 1.55 | 0.43 | 3.33 | 3.06 | 1.82 | 0.28 | 0.00 |

We wish to make our publications widely accessible. Please contact us at publications@ofqual.gov.uk if you have any specific accessibility requirements.

This publication is available at www.gov.uk/ofqual.

Any enquiries regarding this publication should be sent to us at:

Office of Qualifications and Examinations Regulation
Spring Place
Coventry Business Park
Herald Avenue
Coventry CV5 6UB

Telephone  0300 303 3344
Textphone  0300 303 3345
Helpline     0300 303 3346