FORM OF REPORT ON EXAMINATIONS

MSt and MPhil in Linguistics, Philology, and Phonetics Examiners' Report for the academic year 2024–2025

Part I

A. STATISTICS

46 candidates were examined this year: 30 for the MPhil (with 10 of these sitting the first year Paper A exam and 20 taking the second-year exams) and 16 for the MSt. The second year MPhil examinees and 3 of the MSt (Research Preparation) candidates submitted a thesis. At the final examiners' meeting on 8 July 2025, award decisions were made on 29 MPhil candidates (including all first-year MPhil candidates) and on 15 MSt candidates; several of these had extended submission deadlines but could nevertheless be considered because of the hard work and commitment of the assessors of the respective submissions.

Award decisions on the remaining candidates who had even longer extended submission deadlines were made at an additional examiners' meeting on 2 October 2025. The latter meeting took place online, again with the full Board of Examiners present, and with Mrs Camilla Rock (Academic Administrator) in attendance.

Classification statistics for this year and the past two years are given in the following table. At the final examiners' meeting the examiners considered this year's outcome in comparison with that of recent years and noted that it appeared to be broadly in line with these.

(1) Numbers and percentages in each class/category

Unclassified Examinations

Category	Number			Percentage %		
	2024/25	2023/24	2022/23	2024/25	2023/24	2022/23
Distinction	8	9	3	22%	28%	9%
Merit	10	11	8	28%	34%	24%
Pass	16	9	15	43%	28%	45%
Fail	2	3	5	8%	15%	0%

The following units of assessment were examined by a written three-hour examination:

- General Paper: Linguistic Theory

- B(i) Phonetics and Phonology
- B(iv) Historical and Comparative Linguistics
- B(vi) History and Structure of a Language – Breton
- B(vi) History and Structure of a Language: Modern Greek (written)
- B(vi) History and Structure of Latin
- B(vi) History and Structure of Romanian
- B(x) Morphology
- C(i) The comparative grammar of Anatolian and Old Irish
- C(i) The comparative grammar of Greek and Anatolian (Hittite)
- C(ii) The historical grammar of Anatolian and Old Irish
- C(ii) The historical grammar of Greek and Anatolian (Hittite)
- C(iii) Translation from, and linguistic comment upon, texts in Anatolian and Old Irish
- C(iii) Translation from, and linguistic comment upon, texts in Greek and Anatolian (Hittite)
- D(i) The history of French
- D(ii) The structure of French

The following units of assessment were examined by a written submission:

- B(ii) Syntax
- B(iii) Semantics and Pragmatics
- B(v) Psycholinguistics and Neurolinguistics
- B(vi) History and Structure of a Language - French
- B(vi) History and Structure of a Language - Spanish
- B(vii) Experimental Phonetics
- B(viii) Sociolinguistics
- B(ix) Computational Linguistics
- B(x) Attitudes towards Minority Languages in France
- B(x) Digital Scholarship – Methods Paper
- B(x) Lexicography
- B(x) Philosophy of Language
- D(iii) A project on an aspect of the structure or history of French

The following MSt (Research Preparation) and MPhil theses were examined:

- Phonetic realizations of personae in East Asian-American performers' speech (MSt RP)
- A study of subject and object doubling in contemporary, spoken Venetan (MSt RP)
- Wh-movement in Abazgi (MSt RP)
- Clitic placement with finite verbs in Old Occitan
- The development of preverbal negation in Portuguese-based creole in Jakarta
- An Exploration of Sri Lankan English within the Context of Lexicography
- Information Theoretic Approaches to the Ancient Greek Syllabic Scripts
- Intonation of Hong Kong English function words
- Extended Metaphor: Metaphors in Discourse
- Vowels in Canton-Zhanmi: System and Diachronic Change
- Agreement patterns of the verb "to have" in Breton
- Exploring the mechanisms of plausibility illusion in sentence comprehension
- Metrical Phonology in Old and Neo-Hittite
- Temporal Iconicity in Mandarin Narrative Processing

- Unsex Me Here: A Semantic Analysis of Gender-Inclusive Language
- Promissives in Mandarin Chinese
- Analyzing the Foundations of Standardization of the Transcript in United States Court Reporting
- Diachrony of Chinese Resultative Constructions: An Argument Structure Perspective
- Discourse structure of truth and reconciliation commissions
- A Multimodal Analysis of Female Athletes' Representations in Sports Media
- Parallelism in lexical pre-activation and prediction

Use of vivas

On the day of the final examiners' meeting (8 July 2025), no viva was held. One viva was held at the second exam board on Thursday 2 October 2025. This viva examination did not provide evidence that changed the candidate's mark.

Marking of scripts

All scripts were double-marked. Scaling was not employed. The external examiner sampled a wide range of scripts and submissions and paid particular attention to all cases where there was more substantial disagreement between initial marks, where the internal markers had not awarded a pass mark, or where overall classifications were on a borderline. All markers' comments were made available to the External Examiner.

B. NEW EXAMINING METHODS AND PROCEDURES

There were no new methods or procedures operated for the first time in the current academic year.

C. Please list any changes in examining methods, procedures and examination conventions which the examiners would wish the faculty/department and the divisional board to consider.

The examiners would like to encourage the Faculty to consider providing greater clarity on the use of marking scales, particularly at the top and bottom ends of the *de facto* scale. This could help to address anomalies that arise in averaging of marks. This is explained further below.

The examiners are concerned about the handling of plagiarism allegations by the Proctors, and they urge further attention to this.

The examiners recommend clarification on the guidelines for vivas in the MSt, where an overall pass is based on the average mark across all components of the exam.

D. Examination conventions

Candidates are made aware of the examination conventions to be followed by the examiners by means of an email sent to them, to which the conventions were attached. A copy of those conventions is attached. Exam conventions are also made available via the Graduate Linguistics Overview tile on Canvas and referenced in the Graduate Studies Handbook.

Part II

Examiners are asked to ensure that any comments that they do not wish to have transmitted to students are indicated clearly and are kept within the separate *Section E* of this report. The report should include the following sections:

A. GENERAL COMMENTS ON THE EXAMINATION

The examiners were pleased with the overall standard of performance, but they expressed a concern about the relation between marks in Paper A (which most students take) and marks in other components of the examination. This is addressed further below. They also had concerns and/or suggestions about three cross-cutting themes: extensions, marking scales, and AI usage.

Extensions. The widespread use of 1-week self-certified extensions continues to create challenges for the process. In recent years this has become a sector norm in the UK, motivated in part by a desire to not overburden the NHS. The University is well aware of the steep rise in numbers of self-certified extensions, but it shows limited appetite for departing from sector norms by limiting their use. In the current cycle, more than 40% of submissions were delayed by an extension, mostly self-certified. It is tempting to respond to this by simply moving the due date one week earlier, but at present that would penalize the slight majority of students who submit on time.

However, it was noted that the rates of extensions differed between sub-fields. The rate was lower in Syntax and Semantics (25%), and higher in Psycholinguistics, Sociolinguistics, and Experimental Phonetics (65%). This divergence warrants further scrutiny. It is probably not due to the demands of data collection, since that is not required for the Psycholinguistics option. It could reflect students' confidence in their readiness to meet the expectations for each option. The expectations for the Syntax submission are different than for the Semantics submission in terms of the student's need to identify original questions. But both of those sub-fields are covered in the Foundation Course and are a focus of Paper A.

Mark scales and weightings. Under current regulations, the marks for all components are simply averaged. This means that all components are equally weighted in the overall score, despite the fact that some perceive an MPhil thesis to constitute a more substantial piece of work than what is required to pass an option paper. A couple of additional rules govern the criteria for passes and higher categories. In particular, the mark for the thesis places a ceiling on the overall mark, e.g., no Merit overall without a Merit mark on the thesis.

This can create situations where small differences in marks for one option paper can have a dramatic impact on the overall outcome. This year, the examiners faced a clear example of this. If an option paper received a mark of 49, the student would have failed the degree overall. If that same paper received a mark of 50, the student would have passed. And if the paper received a mark of 51, the student would have received a Merit.

The issue of overall weighting was recently reviewed by the Faculty, leading to no change. A key consideration in this outcome was that increasing the weighting for the thesis would not, in recent experience, lead to an overall improvement in results. In that regard, it would not be beneficial for students. Nevertheless, this is independent of the issue of whether rewards align with what is valued.

But there are more pervasive issues. Consider a specific case that arose this year. Other similar examples arise frequently. One candidate had clearly memorized answers to some past questions, hoping that similar questions would arise this year. In a couple of instances they got lucky, in others they did not. When that happened, they wrote the rehearsed answer anyway, even though it bore no relation to the questions in the exam. If this individual had written nothing at all, the markers would likely have given a score of 0. But when faced with

some pages of linguistically coherent but irrelevant material, they felt compelled to assign a mark towards the lower end of their de facto scale, typically in the 35-40 range. The markers of the individual papers, of course, had no way of knowing that they were seeing one piece of a pattern of responses, so they may have felt inclined to give the candidate the benefit of the doubt.

Neither the 0 nor the 40 seems fitting in this instance. A mark of 0 on one question or one submission is almost impossible to recover from. A mark of 40 for entirely irrelevant material assigns substantial additional credit for an answer that has the same level of relevant content as a blank answer.

This maybe reflects a broader issue. The current averaging procedure is based on the assumption that marks are being assigned on a ratio scale. But in practice the scale is not even an interval scale. Rather, current practice is, in effect, an ordinal scale. Averaging has the intended effects for ratio scales and interval scales, but not for ordinal scales. The mark scale purports to be a 0-100 ratio scale, but in practice it is more like a 35-80 ordinal scale. Colleagues rarely assign marks outside those ranges, including for answers that range from "This is 100% irrelevant!" to "This is as good as it gets!"

The Faculty could consider taking steps to avoid unintended consequences of averaging applied to an ordinal scale.

The greatest current impacts arise at the lower end of the scale. If a blank answer is given a mark of 0 and a largely or entirely irrelevant answer is given a mark of 30-40, this means that a great deal of weight is assigned to a difference that is rather small in academic terms. It should not be difficult to clarify guidelines in order to address this unintended consequence.

At the higher end of the scale, decisions about where the de facto upper end of the mark range has consequences for what is valued. The higher the de facto limit of the scale, the more opportunity is created for outstanding work to cancel out shortcomings in other assessments.

AI Usage. Some candidates for this year's submissions included declarations that they had used generative AI tools, e.g., *Claude*, to proofread their work. These individuals are to be commended for their transparency. Several markers had the impression that AI usage was more widespread than the disclosures, and more widespread than in previous years.

Students are already using AI tools widely for a variety of uses, from summarising to coding. The university is now offering free GenAI licenses to all members of the university. There is no question that AI usage will continue to grow and become more normalized.

There is an urgent need for greater clarity on how AI usage should be considered in assessments. This is especially relevant in a Faculty where most PGT assessments are by submission.

We do not have clear answers at present, but it is perhaps helpful to distinguish things that GenAI can and cannot do, and how they might relate to skills that we value.

First, GenAI is pretty good at turning rough-at-the-edges prose into fluent and grammatical English. This is clearly beneficial for non-native speakers of English. It may also be beneficial for those who are tasked with reading the work. Opinions may vary on the importance of these skills to a linguistics assessment.

Second, GenAI is getting fairly good at producing fluent, coherent-sounding summaries of past work. This includes the ability to write well-crafted introductory passages that can place a marker in a positive frame of mind. It is less clear how good these summaries really are,

but they read well. This year's submissions contained several of these. So, some of the things that are expected for a Pass mark are now in reach of GenAI.

Third, GenAI is not particularly good at coming up with original ideas. It is unclear how good it is at giving a good account of why specific ideas or generalizations or hypotheses matter. So, the kinds of things that are typically needed for Merit or Distinction marks are likely out of reach for GenAI.

Some steps that could be taken include the following:

Students could be required to declare how they used AI, if indeed they did so. This might increase transparency, but there is no guarantee of reliable information.

The Faculty could conduct an exercise, possibly in collaboration with students, to identify how AI performs on topics of interest. This could be instructive, for students and staff alike, as a basis for constructive further discussion.

The Faculty could require oral presentations. This is unlikely to be universally welcomed.

The Faculty could move to more unseen exams (and it could provide students with more training). This is unlikely to be broadly popular, and it may lead to greater tension between teaching and research.

It is clear that markers should not be expected to be able to reliably detect AI usage. Markers should be able to focus on the academic content of the work in front of them. Equally, students who do not use AI should not be placed at a disadvantage.

B. EQUALITY AND DIVERSITY ISSUES AND BREAKDOWN OF THE RESULTS BY GENDER

There were no clear trends in relation to gender and the distribution of overall marks. Given the relatively small sample sizes, only rather large discrepancies would be visible.

	Number (M)				Percentage (M)			
Category	2024/25	2023/24	2022/23	2021/22	2024/25	2023/24	2022/23	2021/22
Distinction	4	2	1	1	31%	14%	6%	5%
Merit	3	5	3	5	23%	36%	19%	23%
Pass	5	6	6	4	38%	43%	38%	18%
Fail	1	1	4	0	8%	7%	25%	0%

	Number (F)				Percentage (F)			
Category	2024/25	2023/24	2022/23	2021/22	2024/25	2023/24	2022/23	2021/22
Distinction	4	7	2	5	16%	41%	11%	23%
Merit	8	6	5	4	32%	35%	28%	18%
Pass	12	3	9	3	48%	18%	50%	14%
Fail	1	1	1	0	4%	6%	6%	0%

C. DETAILED NUMBERS ON CANDIDATES' PERFORMANCE IN EACH PART OF THE EXAMINATION

To protect candidate anonymity, this section is included in the confidential section of this report.

D. COMMENTS ON PAPERS AND INDIVIDUAL QUESTIONS

Paper A. The examiners noted that marks on Paper A are lower, on average, than on other assessments for the MSt and MPhil. Overall, many candidates earned distinctions, but only 1 candidate of 24 reached a distinction score on Paper A. This appears to be consistent with patterns from recent years. This raises questions about whether Paper A is an adequate reflection of the aims of the Foundation Course. The examiners recommend attention to this discrepancy.

There are multiple possible reasons for the lower marks on Paper A than on other assessments. (i) The markers may set higher expectations than is reasonable. (ii) Students may receive insufficient training in how to translate successful Foundation Course learning into strong answers in Paper A. E.g., insufficient practice in timed, hand-written exams. (iii) The Paper A questions may be a poor measure of students' learning. (iv) Marks in submitted assessments may be more generous than is warranted. The examiners take no position on whether any of these possibilities is true, but the Faculty is encouraged to take a closer look at this central piece of the PGT curriculum. The Foundation Course has laudable goals, but these goals may be compromised if students perceive Paper A to be a process that typically lowers their overall marks.

It should also be noted that this year saw 0 of 8 candidates achieve a distinction in Bii Phonetics & Phonology, which like Paper A involves a sit-down exam. It is possible that there is an issue with Paper A, or that there is an issue with traditional exam formats for this population.

E. COMMENTS ON THE PERFORMANCE OF IDENTIFIABLE INDIVIDUALS AND OTHER MATERIAL WHICH WOULD USUALLY BE TREATED AS RESERVED BUSINESS

To protect candidate anonymity, this section is included in the confidential section of this report.

F. NAMES OF MEMBERS OF THE BOARD OF EXAMINERS

Professor Willem Hollmann (Edinburgh University, External Examiner)
Professor Colin Phillips (Chair)
Professor Andreas Willi
Professor David Willis

Conclusion

The exam board would like to thank the many people who worked under often tight time constraints to make this year's graduate examinations in Linguistics proceed smoothly. In particular, this would not be possible without the dedication of Camilla Rock and Liberty Braddyll in the Linguistics Academic Office. The exam board encourages the Faculty to pay attention to some structural issues that affect this examination, including (i) the expectations for Paper A, (ii) the way that the marking scale impacts averaging, and (iii) the use of AI in submissions.