

ЭКСПЕРТНЫЙ АНАЛИЗ КРИТЕРИЕВ ОЦЕНКИ ПИСЬМЕННОЙ РЕЧИ

*Работа представлена кафедрой иностранных языков и лингводидактики
СПбГУ.*

Научный руководитель – доктор филологических наук, профессор И. Ю. Павловская

Статья «Экспертный анализ критериев оценки письменной речи» посвящена сравнению значимости критериев в международных тестовых форматах и в российском образовательном контексте. Автор демонстрирует результаты экспертного анализа, который проводился среди преподавателей средней и высшей школы с целью выяснения, какие критерии оценки они применяют при оценивании письменных работ абитуриентов и студентов начальных курсов. Российским экспертам была предоставлена сформированная автором система из 15 критериев оценки, которые должны были быть проранжированы в соответствии с их мнением и опытом с целью выявления общих закономерностей в международных и российской системах оценки.

The article is devoted to the criteria ranging in international standardized tests and Russian educational context. The author outlines the results of the expert analysis which was carried out among the university and school teachers in order to find out which criteria they use while assessing written papers of the school leavers and 1-st and 2-nd year graduate students. The system which included 15 criteria and was developed by the author of the research was given to the Russian experts who were to range the criteria based on their own experience in order to reveal common patterns in international and Russian systems of assessment.

Оценка навыка письменной речи испытуемых, который относится к продуктивным навыкам, представляет собой одну из самых трудоемких задач. Оценка продуктивных навыков всегда субъективна, и основная задача – это достижение максимальной объективности при субъективном оценивании. Поэтому разработка четкой и детализированной шкалы оценки письменной речи представляется необходимой. На ос-

новании фундаментального исследования, проведенного автором данной статьи, было выделено 15 критериев оценки письменной речи, которые применяются в различных европейских и американских шкалах оценки: организация структуры работы, глубина содержания, соответствие заявленной теме, когезия, использование аргументов для подтверждения выдвигаемых тезисов, грамматическая корректность, общее впечатле-

ние от работы, оригинальность суждений, начитанность и информированность по излагаемой теме, наличие примеров, орфографическая корректность, разнообразие вокабуляра, разнообразие языковых структур, пунктуационная корректность, стилистическая корректность.

На основании подробного анализа критериев оценки в таких тестовых форматах как TOEFL, FCE, IELTS, а также, опираясь на рекомендации, представленные в документах ALTE и Совета Европы (CEF), была сформирована система оценки, которая включает не только ранжирование вышеперечисленных критериев, но также и определение веса, который имеет каждый критерий с точки зрения европейских и американских экспертов. (см. табл. 1). Однако, принимая во внимание, что для целей российского образовательного контекста задачей является найти соответствие между критериями, применяемыми как в международной, так и в отечественной практике, был проведен экспертный анализ, о котором и пойдет речь в данной статье.

Цель проведения экспертного анализа – определение отношения отечественных специалистов к критериям оценки письменной речи на английском языке, с тем чтобы сравнить полученные результаты с подходами зарубежных методистов и попытаться найти оптимальный вариант, учитывающий как отечественный, так и зарубежный подход к критериям оценки письменной речи.

Для того чтобы сделать выводы на основе опыта отечественных специалистов, был проведен экспертный анализ, участниками которого являлись преподаватели различных категорий (с точки зрения стажа работы, места работы (средняя школа, высшая школа) и т. д.).

Специалистам в области преподавания иностранных языков был предложен список из 15 критериев оценки письменной речи, который сформировался в результате подробного изучения критериев, применяемых в зарубежных форматах тестов

(TOEFL, FCE, CAE), а также с учетом дескрипторов, применяемых в CEF. Задача преподавателей заключалась в ранжировании критериев (присвоить ранг каждому критерию от 1 до 15 по убыванию – т. е. 1 – наиболее важный критерий при оценке, 15 – абсолютно неважный).

Таким образом, было обработано 50 анкет.

В результате анализа из 15 критериев было выделено 6 (общее впечатление от работы, организация структуры, соответствие заявленной теме, глубина содержания, когезия и грамматическая корректность), которые представляются наиболее важными у различных категорий преподавателей, так как именно у этих критериев медиана ранга менее или равна 5 (см. рис. 1).

Из общей таблицы видно, что в основном эксперты присваивали наивысшие ранги только 6 критериям из 15. К остальным критериям отношение было существенно различным, и по медиане они не вошли в число представленных шести. Так, например, такие критерии, как оригинальность суждений или использование аргументов для подтверждения тезисов или наличие примеров, у одних преподавателей занимали ведущие позиции, в то время как другие сходились во мнении о том, что либо эти критерии нет смысла выделять отдельно либо они не играют важной роли при оценивании письменной работы.

Таблица 1 показывает место критериев в представлениях российских экспертов в сравнении с их местом в списке международных критериев по тем шкалам, которые рассматривались в рамках проведенного исследования.

Из таблицы видно, что критерий «соответствие заявленной теме» оказался наиболее важным для российских экспертов, что в целом соответствует международной практике оценки. Однако в приведенной формулировке, которая наиболее привычна для отечественных методистов, этот критерий обозначается только в оценках тестов Cambridge ESOL и TOEFL. При описа-

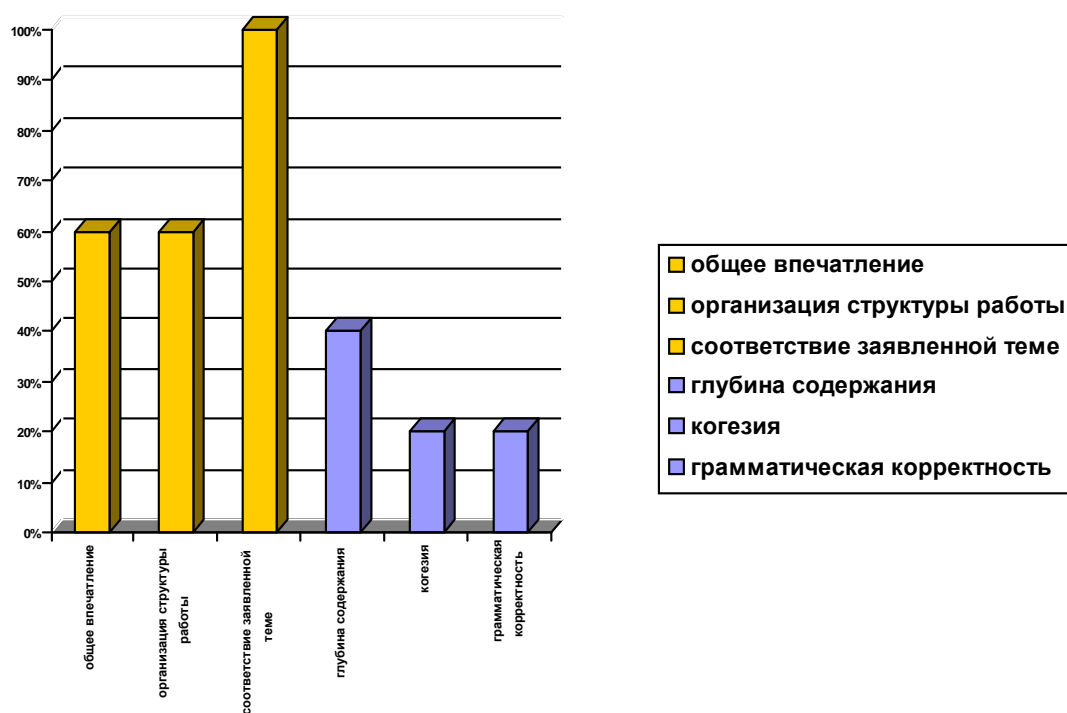


Рис. 1. Распределение рангов у выделенных критериев

Таблица 1

Расстановка приоритетов в критериях оценки письменной речи в различных шкалах

Критерии /шкалы	CESOL	TOEFL	ALTE	CEF	Российские эксперты
Соответствие заявленной теме	2	1	–	–	1
Организация структуры работы	1	2	2	2	2
Когезия	3	3	3	3	5
Грамматическая корректность	4	4	5	4	6
Глубина содержания	–	–	4	5	4
Общее впечатление	–	1	–	–	3
Разнообразие вокабуляра и структур	5	6	6	6	–
Наличие примеров	6	5	–	–	–

нии того, что должен уметь испытуемый на определенном уровне владения письменным языком, в оценках, предлагаемых в CEF и ассоциацией ALTE, этот критерий можно лишь интуитивно выделить из общего описания.

Далее, вторым в списке приоритетов российских экспертов следует критерий «организация структуры работы», и, хотя среди различных групп экспертов место этого критерия варьировалось от 2-го до 5-го, в общей статистике он ставится на второе место, что в целом, соответствует международной практике.

В отношении следующего критерия «общее впечатление», который с небольшим отставанием стоит на третьем месте у российских экспертов в общей статистике, возникают самые серьезные разногласия с международной практикой оценки.

Как видно из табл. 1, в большинстве шкал этот критерий не считается одним из важнейших, за исключением шкалы TOEFL, где в последнее время разработчики ориентируются не столько на аналитический, сколько на холистический подход, т. е. оценка впечатления от работы в целом. Следует заметить, что среди различных

групп российских экспертов этот критерий вызвал самые неоднозначные толкования. Анализ показал, что из 50 респондентов более половины присвоили этому критерию ранг от 1 до 3, в то же время более 10 человек поставили его на последнее место.

В первом случае мы можем говорить о холистическом подходе к оцениванию работы, что, оказалось более свойственно группе «младших преподавателей», во втором случае – проявление аналитического подхода, когда эксперты руководствуются четкими дискретными критериями оценки, из совокупности которых и складывается общее впечатление от работы.

Критерий «глубина содержания», который у российских экспертов занял 4-ю позицию, в международной практике только в шкалах ALTE и European Framework занимает схожую позицию, тогда как в шкалах TOEFL и FCE он даже не входит в число основных критериев. Возможно, разработчики не выделяют этот критерий отдельно, полагая, что важнее раскрыть тему в логической последовательности, а насколько глубока мысль в оцениваемой работе, для выставления высокой оценки не столь важно.

Следует отметить, что в системе дескрипторов в шкалах ALTE и CEF важен факт не просто изложения идеи в соответствующей форме, а то, насколько глубоко испытуемый может мыслить на иностранном языке и излагать свои мысли на письме, что во многом отражает уровень его дискурсивной и общей компетенции.

Творческая составляющая всегда была важна для российского менталитета, именно поэтому такое высокое место данного критерия можно считать вполне обоснованным. Если экзаменатору действительно интересно читать ту или иную письменную работу, другие критерии оценки могут отойти на второй план. В связи с этим данный критерий в сознании российских экспертов может быть близок критерию «общее впечатление от работы». Однако, как показал анализ, критерий «глубина содержания» был более понятен по своей фор-

мулировке и вызвал меньше разногласий у респондентов.

Ранг, присвоенный критерию «когезия» российскими экспертами, наглядно демонстрирует значительное отставание в российской практике оценки одного из важнейших дискурсивных критериев в международной практике. Надо отметить, что все группы респондентов поставили этот критерий на 4-е или 5-е место, что свидетельствует о постепенной переоценке взглядов. Однако 5 человек из числа опрошенных поставили этот критерий на последнее место, не считая необходимым уделять ему должное внимание при оценке или не считая целесообразным выделять его как отдельный критерий.

Удивительным оказался тот факт, что критерий «грамматическая корректность», которому уделялось большое внимание как в оценке, так и при обучении письменной речи, в отечественной практике замыкает «шестерку» основных критериев. Для других шкал, как европейских, так и американских, данный критерий занимает 4-е или 5-е место. Возможно, это происходит в силу того, что грамматическая корректность также относится к лингвистической компетенции, а не к собственно дискурсивной, а в последнее время идет переоценка места дискурсивных критериев.

Таким образом, в результате проведенного исследования удалось сформировать универсальную абстрактную шкалу, состоящую из 15 критериев, тех критериев, которые применяются при оценивании письменных работ в различных тестах. Проведенный анализ показал, что в каждом конкретном случае из общего списка критериев выбираются те или иные в зависимости от целого ряда факторов. Место (ранг) различных критериев также зависит от конкретной ситуации. Тот факт, что российские эксперты выделили 6 критериев, упомянутых выше, не свидетельствует о том, что остальные критерии не принимаются во внимание. Существует множество факторов, которые влияют на набор и приоритетность критериев оценки.

ПЕДАГОГИКА И ПСИХОЛОГИЯ, ТЕОРИЯ И МЕТОДИКА ОБУЧЕНИЯ

Видится целесообразным рассматривать критерии как универсальную сущность, а их отбор и порядок – как конкретную сущность в зависимости от образовательного контекста разных уровней. При этом под образовательным контекстом можно понимать как федеральный уровень, так и университетский уровень или срез по конкретному курсу.

Так как при оценивании письменных

работ обучаемых мы имеем дело с субъективным оцениванием, возникает острая необходимость все критерии и дескрипторы пропускать через определенные процедуры локальной стандартизации в зависимости от ряда факторов: целей обучения, социолингвистических факторов, этнических и национальных особенностей и т. д., что является предметом дальнейшего рассмотрения и будет подробно изложено в следующей статье.