

Die Erprobung landeszentraler Sprechaufgaben in den Abiturprüfungen in Französisch und Englisch an saarländischen Schulen

**Prof. Dr. Claudia Polzin-Haumann und Dr. Christina Reissner,
FR 4.2 Romanistik,
Priv.-Doz. Dr. Stefan Diemer und Marie-Louise Brunner, M.A.,
FR 4.3 Anglistik,
Universität des Saarlandes**

***Wissenschaftliche
Begleitung,
erstellt im Auftrag
des Ministeriums
für Bildung und
Kultur des
Saarlandes***

Inhaltsverzeichnis

1	PROJEKTHINTERGRUND	2
2	DIE STUDIE	2
2.1	RAHMENBEDINGUNGEN UND PRÜFUNGSFORMAT	2
2.2	FORSCHUNGSFRAGEN	3
2.3	VORGEHENSWEISE UND METHODIK	3
3	ERGEBNISSE	4
3.1	SCHULLEITUNGEN	4
3.1.1	ALLGEMEINE ASPEKTE	4
3.1.2	INFORMATION	6
3.1.3	RESONANZ BEZÜGLICH DER ERPROBUNG DER SPRECHPRÜFUNGEN	7
3.1.4	FAZIT DER SCHULLEITUNGEN	8
3.2	LEHRKRÄFTE	10
3.2.1	PRÜFUNGSVORBEREITUNG	10
3.2.2	REAKTIONEN AUF DIE ERPROBUNG DER LANDESZENTRALEN SPRECHPRÜFUNG	13
3.2.3.1	PRÜFUNGSFORMAT	16
3.2.3.2	SCHWIERIGKEITSGRAD	17
3.2.3.3	MATERIALIEN UND INHALTE	20
3.2.3.4	PRÜFUNGSABLAUF	22
3.2.3.5	BEWERTUNG	27
3.2.4	FAZIT DER LEHRKRÄFTE	31
3.3	SCHÜLERINNEN UND SCHÜLER	35
3.3.1	INFORMATION UND ERWARTUNGEN	35
3.3.2	PRÜFUNGSFORMAT UND SCHWIERIGKEIT	39
3.3.3	PRÜFUNGSABLAUF UND BENOTUNG	43
3.3.4	FAZIT UND ANREGUNGEN DER SCHÜLERINNEN UND SCHÜLER	48
3.4	KONSTITUIERUNG DER PRÜFUNGSPAARE	49
4	ZUSAMMENFASSUNG DER ERGEBNISSE UND EMPFEHLUNGEN	51
4.1	WAHRNEHMUNG DURCH DIE AKTEURE	51
4.2	VORBEREITUNG UND ORGANISATION	52
4.3	PRÜFUNGSFORMAT	53
4.4	DURCHFÜHRUNG	54
5	SCHLUSSFOLGERUNGEN UND AUSBLICK	55
6	LITERATURVERZEICHNIS	57

Diese Studie wurde erstellt unter Mitarbeit von Victoria Bohn, Lisa Gaida und Philipp Schwender (Französisch) sowie Sina Burghardt, Elisa Caputo und Annika Haßel (Englisch).

Synthese der Studie:

Die Erprobung landeszentraler Sprechaufgaben in den Abiturprüfungen in Französisch und Englisch an saarländischen Schulen

Im Auftrag des Ministeriums für Bildung und Kultur des Saarlandes wurden Praktiken, Einstellungen und Anschauungen der verschiedenen beteiligten Akteure im Rahmen der Erprobung landeszentraler Sprechaufgaben in den Abiturprüfungen in Französisch und Englisch untersucht. Die Erhebungen erfolgten mittels Fragebögen und leitfadengestützten Interviews und bezogen alle an der Durchführung beteiligten Gruppen ein, um ein möglichst differenziertes Gesamtbild zu entwickeln.

Die Untersuchung bietet als exemplarische Momentaufnahme ohne Anspruch auf Repräsentativität wichtige Anhaltspunkte für eine Gesamteinschätzung der Akzeptanz und Realisierbarkeit landeszentraler Sprechprüfungen im Abitur und zeigt Potentiale für die Fortentwicklung des Fremdsprachenunterrichts Französisch und Englisch im Saarland auf.

1 Projekthintergrund

Im Rahmen einer Neugestaltung der Abiturprüfungen in den Fächern Französisch und Englisch wird erwogen, in beiden Fächern eine mündliche Prüfungskomponente in Form von landeszentralen Sprechprüfungen einzuführen. Dafür sollen Prüfungen und Bewertungsraster in den beiden Kern-Fremdsprachen des saarländischen Sprachenkonzepts gemeinsam vorbereitet und vergleichbar gestaltet werden.

Zur Vorbereitung der projektierten erstmaligen Durchführung der mündlichen Abiturprüfungen (2018) fand auf Initiative des Ministeriums für Bildung und Kultur (MBK) im Oktober 2015 eine Erprobung der Sprechprüfungen unter Abiturbedingungen statt. Die wissenschaftliche Begleitung dieses probeweisen Durchlaufs erfolgt durch Frau Prof. Dr. Claudia Polzin-Haumann und Frau Dr. Christina Reissner seitens der Professur für Romanische Sprachwissenschaft (Schwerpunkt Sprachlehrforschung Französisch, FR 4.2 Romanistik) und durch Herrn Priv.-Doz. Dr. Stefan Diemer und Frau Marie-Louise Brunner, M.A., seitens der Professur für Anglistische Sprachwissenschaft (FR 4.3 Anglistik, Amerikanistik und Anglophone Kulturen).

Zielsetzung der wissenschaftlichen Begleitung ist die Entwicklung eines möglichst umfassenden Gesamtbilds zu dem neuen Prüfungsformat.

Die wissenschaftliche Begleitung richtet sich auf die Gestaltung und Umsetzung der probeweise durchgeführten Sprechprüfungen. Dabei werden einerseits die tatsächlichen Abläufe, andererseits die Einschätzungen und Einstellungen der beteiligten Akteure in den Blick genommen. Ziel der Studie ist es, die Abläufe sowie die Akzeptanz der Prüfungen bei den Akteuren zu untersuchen, mögliche Probleme im Vorfeld der Einführung zu identifizieren sowie daraus Möglichkeiten zur Weiterentwicklung abzuleiten.

2 Die Studie

2.1 Rahmenbedingungen und Prüfungsformat

Die Erprobung landeszentraler Sprechaufgaben unter Abiturbedingungen in Französisch und Englisch wurde am 14./15.10.2015 an 44 saarländischen Schulen mit gymnasialer Oberstufe durchgeführt¹. Die Prüfungsaufgaben wurden vom MBK bereitgestellt. Es handelt sich um

¹ Diese Sprechprüfungen ersetzen im Schuljahr 2015/16 an den teilnehmenden Schulen die mündliche Kursarbeit, die im 3. Kurshalbjahr der Hauptphase der gymnasialen Oberstufe vorgesehen ist.

Paarprüfungen mit monologischen und dialogischen Anteilen (bei einer ungeraden Zahl an Prüflingen in einem Kurs findet eine Dreierprüfung statt). Die Prüfungsgruppen werden durch Losverfahren konstituiert².

Zu Beginn der Prüfung ist eine kurze Aufwärmphase vorgesehen. Für den monologischen Teil erhält jeder Prüfling einen Sprechimpuls, der in ein bis zwei Minuten zu erfassen ist. Im dialogischen Teil setzen sich die Prüflinge im Gespräch weiter mit dem Impuls auseinander.

Die Prüfungsdauer beträgt 15-20 Minuten (Paarprüfung) bzw. 23-30 Minuten (Dreierprüfung); für die Notenfindung sind 10 Minuten (bzw. 15 Minuten bei der Dreierprüfung) vorgesehen. Die Prüfungskommission besteht aus zwei Lehrkräften pro Kurs (die Kurslehrkraft und ein/e Zweitprüfer/in). Beide Lehrkräfte führen Protokoll. Die Protokollbögen und Bewertungsraster sind für die beiden Fremdsprachen unterschiedlich.

Im Vorfeld der Erprobung (Juni, Juli, September 2015) fanden ganztägige Fortbildungen statt, um die Lehrkräfte mit dem Prüfungsformat und den Abläufen vertraut zu machen.

Die landeszentralen Sprechprüfungen wurden insgesamt in 133 Kursen (99 Englisch, 34 Französisch) mit 2273 Schülerinnen und Schülern erprobt.

2.2 Forschungsfragen

Die Erhebung nimmt verschiedene Parameter in den Blick, die für eine mögliche Einführung landeszentraler Sprechaufgaben in den Abiturprüfungen in Französisch und Englisch an saarländischen Schulen mit gymnasialer Oberstufe von Relevanz sind. Im Fokus stehen dabei systematisch alle an der Erprobung im Oktober 2015 beteiligten Akteure: die Schulleitungen und Lehrkräfte sowie die Schülerinnen und Schüler.

Das Erkenntnisinteresse richtet sich u.a. auf folgende Bereiche:

- Organisatorische und inhaltliche Vorbereitung der Prüfungen;
- Wahrnehmung der Prüfung durch die verschiedenen Akteure (Schulleitungen, prüfende und organisierende Lehrkräfte, Schülerinnen und Schüler);
- Prüfungsformat und erfasste Kompetenzen der Schülerinnen und Schüler;
- Gestaltung der Sprechaufgaben, insbesondere im Hinblick auf Prüfungsgegenstände, Materialien, Inhalte, Prüfungsablauf und Benotung;
- Praktische Aspekte bei der versuchsweisen Durchführung der Prüfungen.

Aus der Summe und der Korrelation der erhobenen Daten sollen mögliche Impulse für die Weiterentwicklung des Prüfungsformats abgeleitet werden.

2.3 Vorgehensweise und Methodik

Im Anschluss an die probeweise Durchführung der landeszentralen Sprechprüfungen erfolgte eine fragebogenbasierte Erhebung (geschlossene/ teilgeschlossene/ offene Fragen) an allen teilnehmenden Schulen, der sich leitfadengestützte Interviews an ausgewählten Standorten anschlossen.

Die Datenerhebung wurde anonym und unter Einhaltung der rechtlichen Vorgaben seitens des MBK sowie des Datenschutzrechts und in Abstimmung mit den beteiligten Akteuren

² Auf Wunsch des MBK haben zwei Schulen an der Erprobung teilgenommen, ohne das Losverfahren anzuwenden. Die hierzu vorliegenden Rückmeldungen wurden separat ausgewertet und – unter Berücksichtigung ihres deutlich geringeren Umfangs – im Hinblick auf die einschlägigen Fragen zu den übrigen Ergebnissen in Bezug gesetzt; vgl. Kap.3.4).

durchgeführt. Die Konzeption und Durchführung der Studie erfolgte durch die Professuren für Romanische Sprachwissenschaft und Anglistische Sprachwissenschaft der Universität des Saarlandes in Abstimmung mit dem MBK.

Allen teilnehmenden Schulen wurden über das MBK zeitgleich mit den Prüfungsaufgaben die Fragebögen zugesandt, die sich jeweils an die Schulleitungen, die prüfenden Lehrkräfte und die geprüften Schülerinnen und Schüler richteten. Somit wurden alle Akteure flächendeckend angesprochen.

Der Gesamttrücklauf gestaltet sich folgendermaßen:

Akteure	Englisch	Französisch
Lehrkräfte	140	55
Schülerinnen und Schüler	1239	391
Schulleitungen	38	

Die Auswertung kombiniert quantitative und qualitative Methoden. Die eingegangenen Fragebögen wurden zunächst eindimensional statistisch ausgewertet. Auf der Grundlage der Zwischenergebnisse wurden im November und Dezember 2015 an sieben ausgewählten Schulen vertiefende Interviews mit Schulleitungen, Oberstufenkoordinationen und Lehrkräften geführt, um ein qualitativ differenziertes Bild zu erhalten. Dabei wurden Gespräche mit 12 Mitgliedern erweiterter Schulleitungen sowie mit 19 Englisch- bzw. Französischlehrkräften geführt. Die Interviews wurden schriftlich zusammengefasst und ausgewertet.

Die einzelnen Auswertungen wurden zueinander in Bezug gesetzt. Diese Bezüge wurden sowohl zwischen den unterschiedlichen Erhebungsarten (schriftlich – mündlich) als auch zwischen verschiedenen Akteuren (z.B. Lehrkräfte – Schülerinnen und Schüler) und schließlich bezüglich einzelner Fragenkomplexe hergestellt. Ziel war es, Gemeinsamkeiten (übereinstimmende/ähnliche Sichtweisen, Einschätzungen, Wahrnehmungen; verstärkte Tendenzen) oder Unterschiede (unterschiedliche Sichtweisen, Einschätzungen, Wahrnehmungen; ggf. unterschiedliche Äußerungen bzw. Nuancierungen von Positionen derselben Akteursgruppe in der schriftlichen oder der mündlichen Befragung) zu ermitteln.

Die Evaluierung berücksichtigt sowohl die fachliche Differenzierung der beiden Fremdsprachen Französisch und Englisch als auch übergreifende Aspekte, die sich durch das Format der Sprechprüfungen sowie die Vorgaben der Bildungsstandards und der Kompetenzorientierung ergeben.

3 Ergebnisse

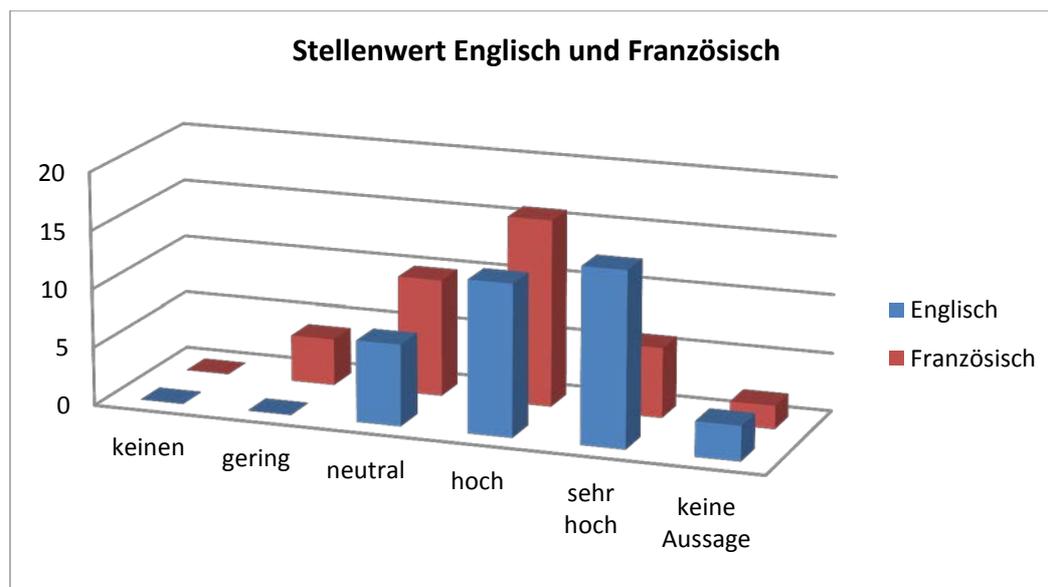
3.1 Schulleitungen

Die folgenden Ausführungen stützen sich auf die Auswertung der Fragebögen sowie die leitfadengestützten Interviews mit den Mitgliedern der erweiterten Schulleitungen an ausgewählten Standorten.

3.1.1 Allgemeine Aspekte

Alle rückmeldenden Schulen fördern den Fremdsprachenunterricht auf vielfältige Art und Weise, darunter Schulpartnerschaften, Teilnahme an Austausch- und Begegnungsprojekten, Fremdsprachenassistenten, im Rahmen von bilingualen Sachfachunterricht oder ABIBAC. Sowohl interne als auch externe Evaluationen werden genutzt, neben Vergleichsarbeiten z.B. auch Zertifikate (DELTA, Cambridge) oder VERA 8. Allerdings wird der Stellenwert von Englisch

und Französisch unterschiedlich eingeschätzt, i.d.R. in Abhängigkeit von der generellen fremdsprachlichen Orientierung der jeweiligen Schule:



Grafik 1: Schulleitung - Stellenwert Englisch (blau) und Französisch (rot)

Die Frage nach der persönlichen Einschätzung der landeszentralen Erprobung der Sprechprüfungen ergibt ein differenziertes Bild und führt unmittelbar zu einigen der besonders polarisierenden Aspekte.

Einerseits wird die Bedeutung der Sprechfertigkeit als Teil einer umfassenden fremdsprachlichen Kompetenz erkannt; der mit der Prüfung einher gehenden Stärkung der Mündlichkeit im Unterricht wird ein motivierender Effekt zugeschrieben. Die Prüfungen werden in den Freitexten als „lebensnah und schülergerecht“, außerdem „gut vorbereitet und prüfungstechnisch ausgereift“ oder als „sinnvolle Kompetenzüberprüfung“ beschrieben.

Andererseits werden häufig (von 55,3% der Schulleitungen) der hohe Organisationsaufwand, die starke Belastung der prüfenden Lehrkräfte (18,4%) und der gedrängte Zeitrahmen (10,5%) bemängelt. Als weiterer negativer Aspekt wird der Unterrichtsausfall genannt, der aus Sicht einiger Schulleitungen für den reibungslosen Ablauf der landeszentralen Sprechprüfungen unvermeidbar ist, und der bisweilen mit dem Plädoyer verbunden wird, die Sprechprüfungen nicht in die Abiturprüfungen aufzunehmen, sondern als „normales Klausurformat“ zu belassen.

Die Gesamteinschätzung hängt im Einzelnen von der individuellen Gewichtung der genannten Parameter ab. Einschätzungen, die Organisationsaufwand und Ertrag als „absolut im Verhältnis“ ansehen, stehen Aussagen gegenüber, nach denen die Gesamtorganisation als zu aufwändig eingestuft wird, zumal wenn die Prüfungen zusätzlich zur Hörverstehensprüfung³ durchgeführt werden müssten (vgl. auch 3.1.4).

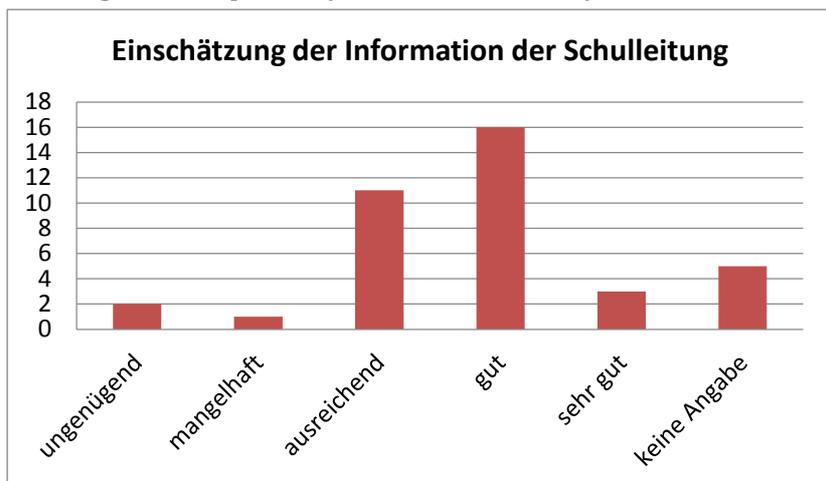
Die im Vorfeld der Prüfungen angebotenen Fortbildungen und die Prüfungsmaterialien werden insgesamt positiv beurteilt (z.B. „zentrale Aufgabenstellung ist große Entlastung“).

Einzelne Schulleitungen weisen auf Unterschiede zwischen den Französisch- und Englischprüfungen hin; hier werden etwa die unterschiedlich gestalteten Protokollbögen, eine unterschiedliche Vorbereitungszeit und die teils nicht vergleichbare Komplexität der Aufgaben genannt.

³ Die Hörverstehensprüfung wird in unterschiedlicher Weise thematisiert: Für einige ist sie wesentlich komplizierter zu organisieren als die Sprechprüfung; andere stufen ihre Organisation als einfacher ein, da sie für alle Prüflinge gleichzeitig angesetzt wird.

3.1.2 Information

Die Information der Schulleitungen zu den Prüfungen durch das MBK wird insgesamt überwiegend eher positiv (im Mittel 3,52 von 5) bewertet:



Grafik 2: Schulleitung - Einschätzung der Information der Schulleitung

Kritisch beurteilt wird allerdings in der schriftlichen Befragung wie in den Interviews sowohl von den Schulleitungen als auch insbesondere von den Oberstufenkoordinatoren die zu kurzfristige Information der Beteiligten seitens des Ministeriums zu den konkreten Abläufen und Aufgabenstellungen. Die fehlende Einbeziehung der Oberstufenkoordinatoren im Vorfeld der Prüfungen wird mehrfach bemängelt.

Die Vorbereitung der Lehrkräfte wird aus Sicht der Schulleitungen ebenfalls meist mit *gut* (im Mittel 3,82 von 5) beurteilt. 65,8% der Schulleitungen bewerten die Vorbereitung der Lehrkräfte mit *gut* bis *sehr gut*, allerdings bewerten auch 23,7% sie mit *ausreichend* bis *mangelhaft*.

In diesem Kontext werden die Handreichungen des Ministeriums in Freitextfragen und Interviews oft positiv bewertet, da die Lehrkräfte dadurch entlastet würden. Die Handreichungen seien wegen der Komplexität der Aufgabenstellung „sehr willkommen“; es wird auch der Wunsch geäußert, diese zu ergänzen. Außerdem wird die zentrale Aufgabenstellung als „große Entlastung“ empfunden. Allerdings wird angemerkt, dass eine Integration in den Unterricht wegen der späten Ankündigung nicht im gewünschten Umfang möglich war.

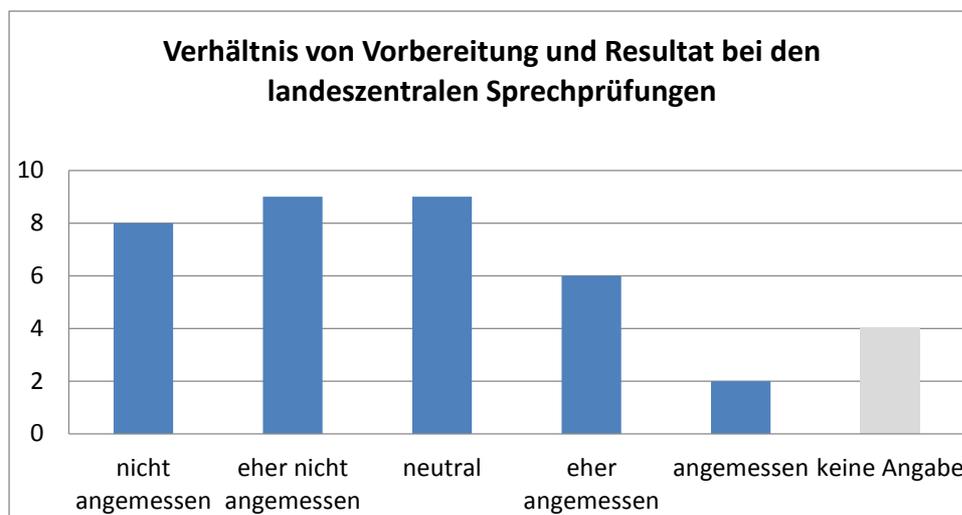
Es wird u. a. angeregt, dass den Lehrkräften eine Vorbereitungszeit mit den konkreten Prüfungsunterlagen ermöglicht werden sollte, z.B. am Tag vorher. Dies korreliert mit Forderungen der Lehrkräfte nach mehr Vorbereitungszeit mit den konkreten Prüfungsunterlagen (vgl. 3.2.3.3).



Grafik 3: Schulleitung - Vorbereitung der Lehrkräfte insgesamt

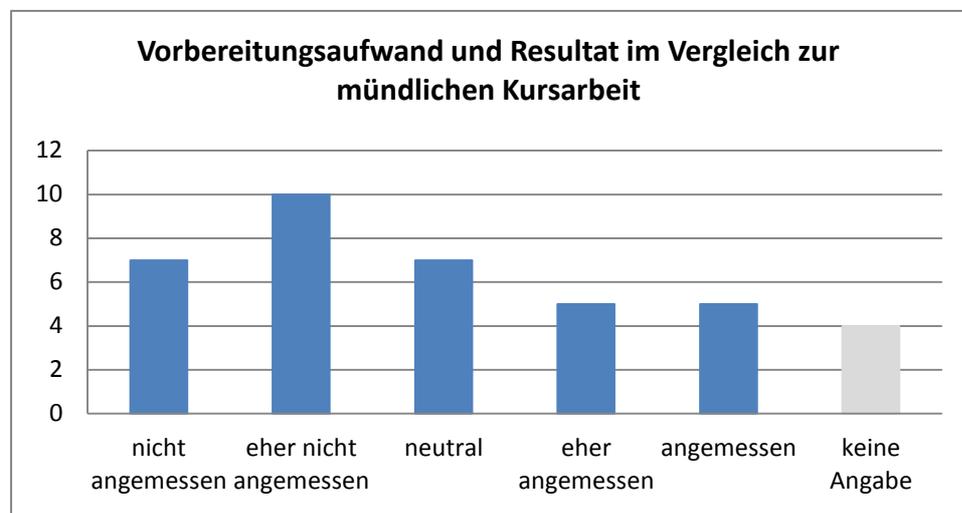
3.1.3 Resonanz bezüglich der Erprobung der Sprechprüfungen

Die Frage nach dem Verhältnis von Vorbereitungsaufwand und Resultat ergibt ein differenziertes Bild (vgl. auch oben 3.1.1). Es ist zu erkennen, dass der Aufwand insgesamt als relativ hoch angesehen wird; das Verhältnis von Vorbereitungsaufwand und Resultat wird zwischen *neutral* und *nicht angemessen* (im Mittel 2,55 von 5) eingeschätzt, wie aus Grafik 4 hervorgeht:



Grafik 4: Schulleitung - Verhältnis von Vorbereitungsaufwand und Resultat bei den landeszentralen Sprechprüfungen

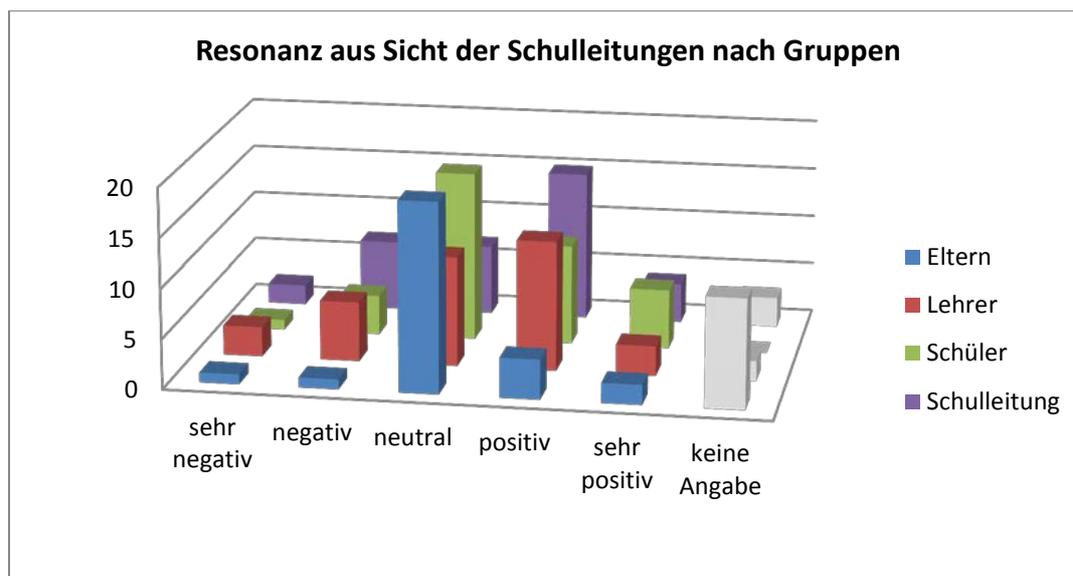
Insbesondere im direkten Vergleich zum Vorbereitungsaufwand für die mündliche Kursarbeit (die schulintern in Verantwortung der Fachlehrkräfte organisiert wird) wird der Aufwand für die landeszentrale Sprechprüfung von beinahe der Hälfte der Schulleitungen als *eher nicht* oder *nicht angemessen* eingeschätzt:



Grafik 5: Schulleitung - Verhältnis von Vorbereitungsaufwand und Resultat im Vergleich zur mündlichen Kursarbeit

Was die Resonanz bei den anderen beteiligten Akteuren angeht, ist diese aus Sicht der Schulleitungen differenziert, aber grundsätzlich positiv. Im Einzelnen wird die Wahrnehmung der landeszentralen Sprechprüfungen durch das Fachkollegium und die Schulleitungen selbst als überwiegend *positiv*, durch die Schülerinnen und Schüler als *neutral* bis *positiv* bewertet. Die Haltung der Eltern wird insgesamt zum Zeitpunkt der Befragung als *neutral* eingeschätzt;

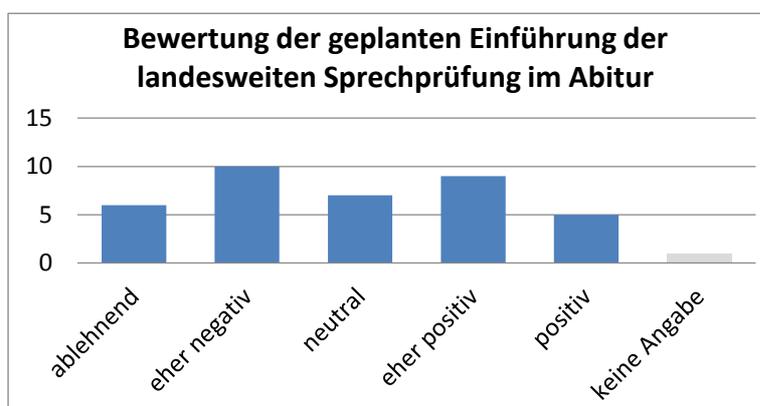
mehrheitlich waren den Schulleitungen keine Elternreaktionen bekannt, lediglich in einem Fall werden zwei Elternbeschwerden vermerkt. Grafik 6 stellt dies zusammenfassend dar:



Grafik 6: Schulleitung - Resonanz an der Schule nach Gruppen

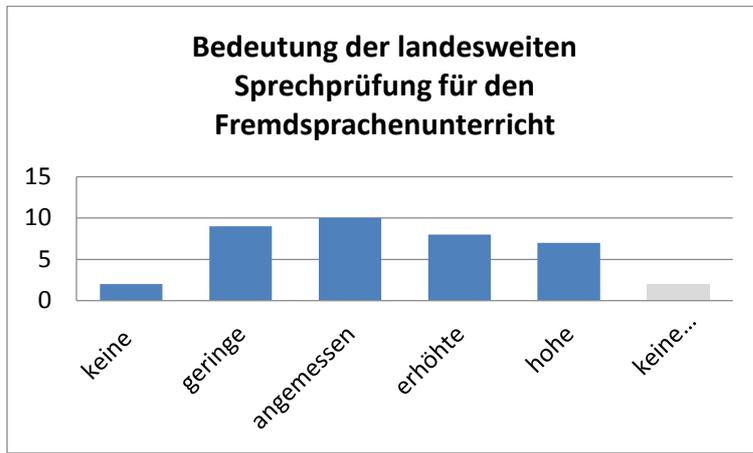
3.1.4 Fazit der Schulleitungen

Die konkrete Frage nach der Haltung der Schulleitungen zu der geplanten Einführung der Sprechprüfungen ergibt ein ambivalentes Bild. 42% der Schulleitungen sind diesbezüglich skeptisch, 37% stehen der Einführung positiv gegenüber und 18,4% vertreten eine neutrale Position.



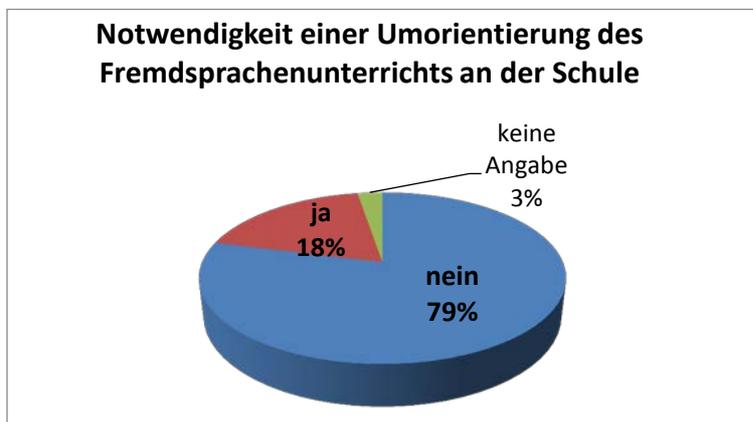
Grafik 7: Schulleitung - Bewertung der geplanten Einführung der landesweiten Sprechprüfung im Abitur

Offenbar werden in der Gesamteinschätzung die Kritikpunkte im organisatorischen Bereich über die inhaltlich-fachlichen Aspekte gestellt. Dies wird deutlich, wenn nach der Bedeutung der Sprechprüfungen für den Fremdsprachenunterricht insgesamt gefragt wird. Grafik 8 visualisiert das gemischte Fazit mit dennoch positiver Grundtendenz (vgl. auch 3.1.1). Insgesamt bewerten 65,8% der Schulleitungen die Bedeutung der Prüfungen als *angemessen* bis *hoch*:



Grafik 8: Schulleitung - Bedeutung der geplanten Einführung der landesweiten Sprechprüfung für den Fremdsprachenunterricht

Aus Sicht der meisten Schulleitungen ist für die Einführung der landeszentralen Sprechprüfungen keine Umorientierung des Fremdsprachenunterrichts an ihrer Schule erforderlich. Einzelne Schulleitungen vermerken die Notwendigkeit „zusätzliche[r] Übungen“, eines höheren Kommunikationsanteils im Unterricht oder der Durchführung von Sprechprüfungen bereits in der Unterstufe.



Grafik 9: Schulleitung - Notwendigkeit einer Umorientierung des Fremdsprachenunterrichts an der Schule

Die Schulleitungen geben aus den Erfahrungen mit dem probeweisen Durchlauf der Sprechprüfungen unter Abiturbedingungen zahlreiche konkrete Anregungen für die weitere Ausgestaltung dieses Formats. Hier werden in erster Linie mehr Zeit für die Notenfindung sowie flexiblere und großzügigere Zeitpläne (ggf. unter Einplanung eines weiteren Prüfungstags) eingefordert; außerdem sollten nach Ansicht mehrerer Schulleitungen durchgehend vergleichbare Aufgaben und einheitliche, weniger komplexe Protokollbögen eingesetzt werden. Schließlich wird insgesamt eine Vereinfachung des Verfahrens mit mehr Eigenverantwortung der Schulen (z.B. im Fall von kranken Prüflingen und der sich daraus ergebenden Notwendigkeit der Reorganisation von Paarungen; Nachprüfungen) und der Fachlehrkräfte (z.B. in Bezug auf das vorherige Einsehen von Prüfungsthemen) als notwendig erachtet.

Während das angewandte Losverfahren in der schriftlichen Befragung nur vereinzelt Erwähnung findet, wird es in den Interviews kontrovers diskutiert: So wird es einerseits mit dem Hinweis auf mögliche „unglückliche Paare“ als „nicht sinnvoll“ angesehen, wobei darauf verwiesen wird, dass die Kompetenz zur Gruppeneinteilung auch bei den Lehrkräften vor Ort vorhanden sei. Andererseits wird das Losverfahren als „das gerechteste“ bzw. „simpelste“ Verfahren eingestuft, mit dem nach anfänglicher Skepsis auch die Schülerinnen und Schüler zufrieden seien (vgl. auch Kap. 3.3.2 und 3.4).

3.2 Lehrkräfte

Das Meinungsbild bei den Lehrkräften gestaltet sich trotz einiger durchgehend kritischer Stimmen insgesamt mehrheitlich eher positiv. Die geäußerte Kritik bezieht sich ganz überwiegend auf organisatorische Detailfragen; inhaltlich findet die geplante landeszentrale Sprechprüfung vornehmlich Zustimmung bei den Lehrenden.

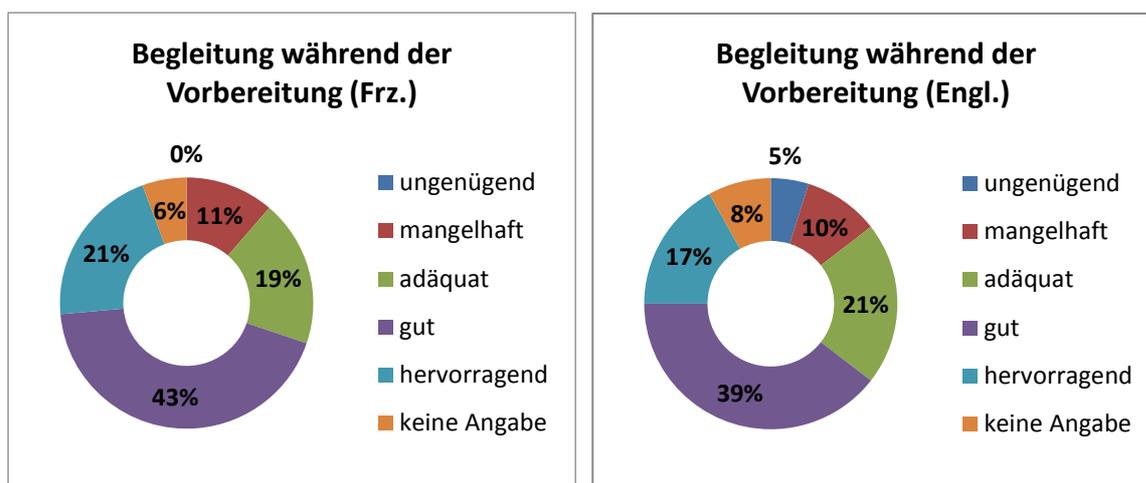
3.2.1 Prüfungsvorbereitung

Die Vorbereitungsangebote des MBK und des LPM werden insgesamt überwiegend positiv beurteilt: 62% der Französisch- und 60% der Englischlehrkräfte bewerten sie mit *adäquat* bis *gut*, 21% der Französisch- und 17% der Englischlehrkräfte sogar als *hervorragend*. Allerdings schätzen beinahe ebenso viele (11% der Französischlehrenden und 15% der Englischlehrenden) die Vorbereitung als *mangelhaft* bzw. *mangelhaft* und *ungenügend* ein.

Die Lehrkräfte fühlten sich also insgesamt *adäquat* bis *hervorragend* vorbereitet, vor allem durch die Fortbildung und die Materialien des LPM. Daneben wurden auch eigene (alte) Materialien genutzt; zudem fand (häufig über Fächergrenzen hinweg) ein reger Austausch zur Vorbereitung statt. Trotzdem wurden in den Fragebögen oft mehr Materialien (Musteraufgaben, Übungen, Übungsvideos) und ausführlichere oder weitere Fortbildungen gewünscht.

Auch in den Interviews wurde das Vorbereitungsformat mit Fortbildung am LPM und einem Materialpaket insgesamt gutgeheißen:

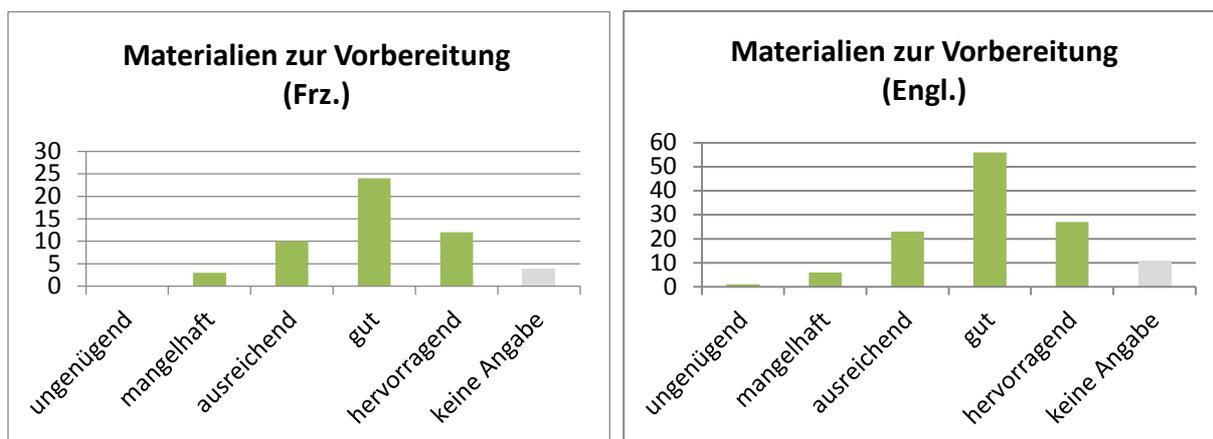
- „Es war eine schöne Erleichterung, dass wir nicht selbst Themen aussuchen mussten und dass ein zeitlicher Rahmen vorgegeben wurde. Die Vorbereitung war dadurch besser.“ (E)
- „[...] das ist schon ein Service! Muss man hier mal noch hervorheben! Das waren viele Aufgaben. Wir sind die alle durchgegangen.“ (E)
- „Die Fortbildung war wirklich sehr hilfreich, vor allem die Mitschnitte von Prüfungen, dadurch hat man ein Gefühl dafür bekommen.“ (E)
- „Wir haben die Fortbildung noch im letzten Schuljahr gemacht. Ich konnte dann mit dem Material im Juli noch arbeiten. Für jemanden, der im September eine Fortbildung gemacht hat, ist es nicht möglich, das zu machen.“ (E)
- „Erleichternd war, dass man die Prüfung nicht selbst aufsetzen musste.“ (F)
- „Materialien waren auch gut.“ (F)



Grafik 10a und 10b: Lehrkräfte - Begleitung während der Vorbereitung

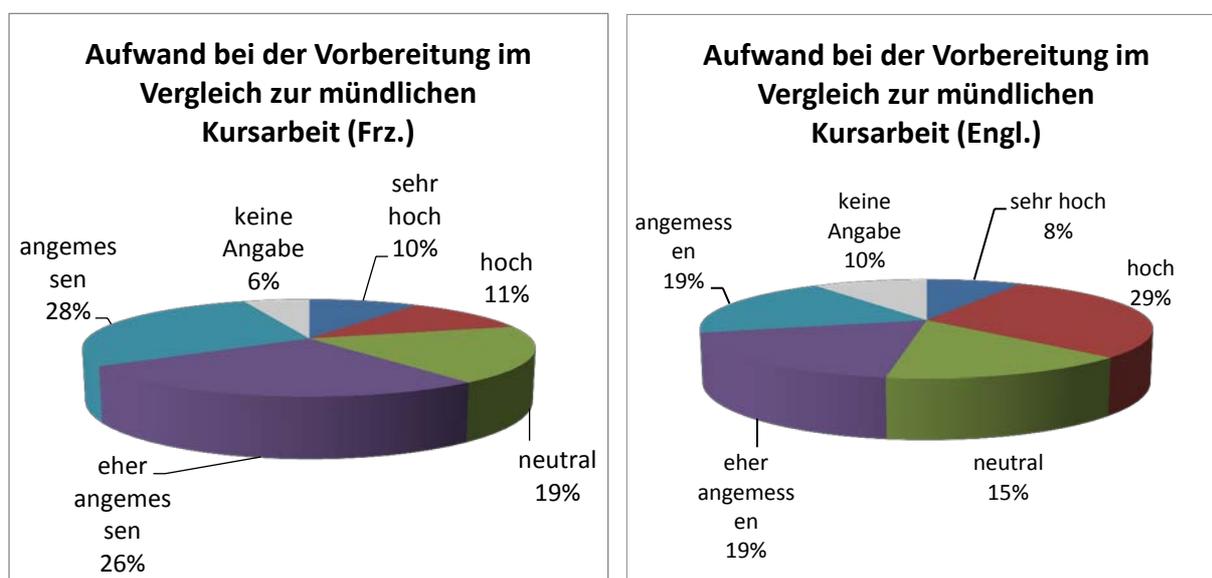
Die zur Verfügung gestellten Vorbereitungsunterlagen wurden von den meisten Lehrenden gelobt – 66,9% der Englischlehrkräfte und 67,9% der Französischlehrkräfte empfanden sie als *gut* oder *hervorragend*. In der Befragung hoben die Lehrkräfte die „gute, ansprechende Qualität“

(E) hervor. Die Materialien seien „übersichtlich und transparent“ (E); es gebe „keine bösen Überraschungen“ (E). Vor allem die klare Struktur, die hervorragende Druckgestaltung (Farbe) und die zentrale Erstellung wurden als positiv empfunden. Kritik wurde hier nur sehr vereinzelt geübt.



Grafik 11a und 11b: Lehrkräfte - Bewertung der Materialien zur Vorbereitung

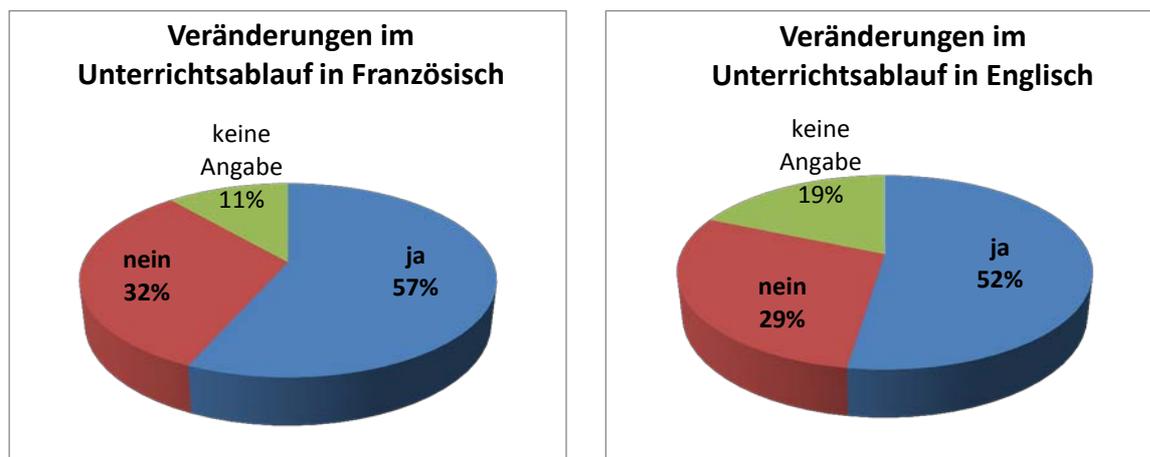
Hinsichtlich des Vorbereitungsaufwandes entstehen nach Ansicht der Lehrkräfte durch die landesweite Sprechprüfung im Fach Französisch keine großen Unterschiede im Vergleich zur mündlichen Kursarbeit, da der Aufbau der landeszentralen Sprechprüfung dem der mündlichen Kursarbeit weitgehend entspreche (vgl. Grafik 12a). Eine Lehrkraft beschreibt diesen Sachverhalt im Interview mit den Worten: „Das Schema war ritualisiert.“ (F) Der Aufwand wird von den Englischlehrkräften im Vergleich höher eingeschätzt (vgl. Grafik 12b); so wird in einem der Interviews beispielsweise angemerkt: „An sich eine ganz tolle Sache, aber der Organisationsaufwand ist viel zu hoch“. (E)



Grafik 12a und 12b: Lehrkräfte - Aufwand bei der Vorbereitung der landeszentralen Sprechprüfung im Vergleich zur mündlichen Kursarbeit

Befragt nach Veränderungen im Unterrichtsablauf gibt die überwiegende Mehrheit der Lehrkräfte aus beiden Fächern an, dass sich die probeweise landeszentrale Sprechprüfung auf den Unterricht auswirke, insbesondere was den Umgang mit den Impulsen und die Vorbereitung auf den konkreten Prüfungsablauf betrifft (vgl. Grafik 13a und 13b).

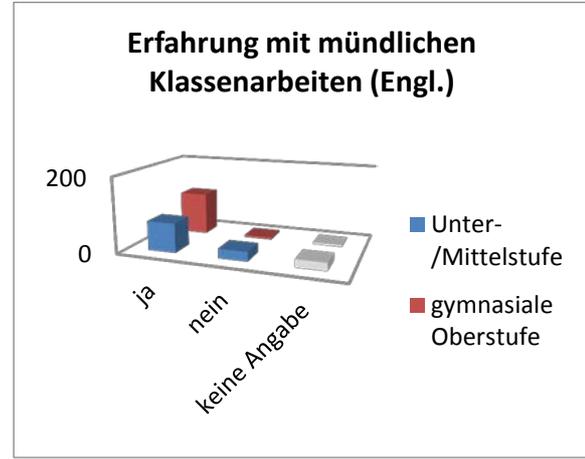
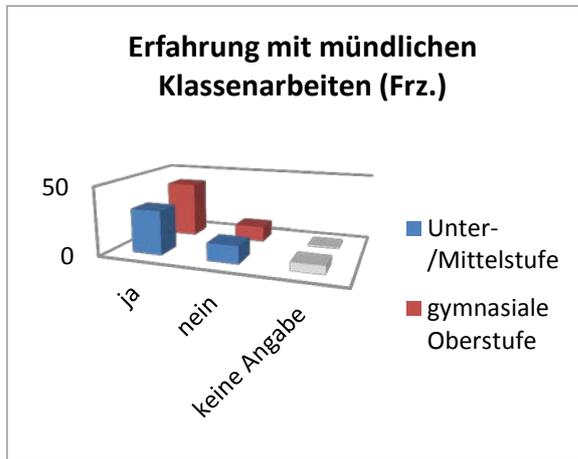
Die Englischlehrenden nennen hier beispielsweise Sprechtraining, Einüben von konkreten Kommunikationsstrategien und des Prüfungsformats (*mock exams*) sowie allgemein mehr Raum für Sprechübungen im Unterricht. Aus den Interviews geht hervor, dass anhand eigener und der Materialien aus der Fortbildung verstärkt Diskussionsanlässe integriert und während des Unterrichts Diskussionen in Kleingruppen geführt wurden, um die Schülerinnen und Schüler auf die Sprechprüfung vorzubereiten. Die „Schülerinnen und Schüler hatten viele Fragen zu Inhalt und Ablauf der Prüfungen und der Prüfungsformate“ (E), wie eine Englischlehrkraft betont. Gleiches gilt für den Bereich Französisch. Die Impulse wurden „richtig intensiv geübt“ (F). Zusammenfassend formuliert eine Französischlehrkraft im Interview: „Jedoch musste dieses Material, wie es hier vorliegt, noch mal gesondert trainiert werden.“ (F) Häufig wurde auch „der Ablauf geübt, [...] das war sehr sinnvoll in Französisch, dieses Schema genau zu üben.“ (F) Eine Englischlehrkraft bemerkt: „Ich hab das [die 3er-Prüfung] auch geübt, für den Fall der Fälle und ich habe auch jede Prüfung im Unterricht gelöst.“ (E) Die befragten Lehrkräfte setzten für die Vorbereitung Zeiträume von drei bis vier Wochen an; eine Englischlehrkraft merkt an, dass die „letzten beiden Wochen allein zum Üben der Sprechfertigkeit genutzt“ (E) wurden.



Grafik 13a und 13b: Lehrkräfte - Veränderungen im Unterrichtsablauf durch die landeszentrale Sprechprüfung

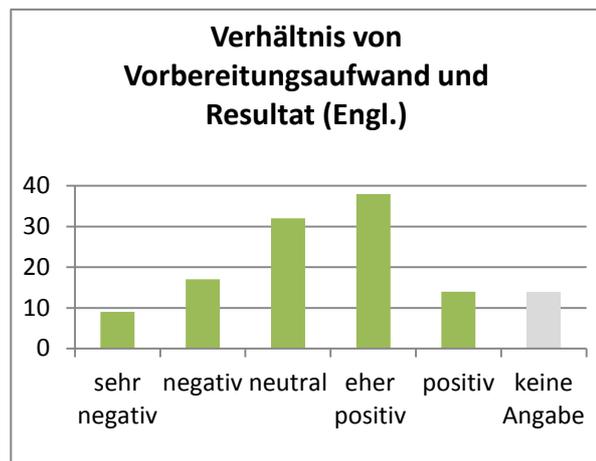
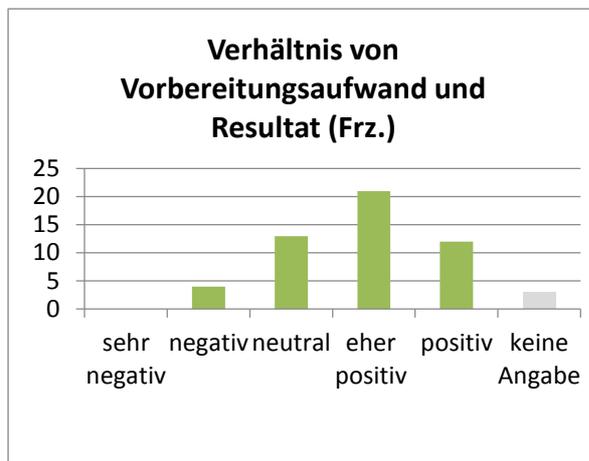
Die Lehrkräfte haben ganz überwiegend Erfahrungen mit der Prüfung des Kompetenzbereichs „Sprechen“, i.d.R. in Form einer mündlichen Kursarbeit, allerdings hauptsächlich in der Oberstufe. Ein Interviewpartner bezeichnet die mündliche Kursarbeit sogar als Schul-„[k]ultur“. (F)

Eine Französischlehrkraft erläutert weiter: „Wir haben ja genug Möglichkeiten, die mündliche Leistung von den Schülerinnen und Schülern im Unterricht zu bewerten und das hat schon an Bedeutung in den letzten Jahren zugenommen. Ich gebe den Schülerinnen und Schülern im Rahmen der Kursarbeit Rückmeldung, wie ich ihre mündliche Leistung einschätze. Dann hat man noch die Möglichkeit der mündlichen Klassenarbeit.“ (F) Eine Englischlehrkraft merkt im direkten Vergleich mit der mündlichen Klassenarbeit an, dass es „[...] besser [war], dass sich die Schülerinnen und Schüler im Vorfeld treffen und gemeinsam (im Unterricht) üben konnten; die Gesprächspartner sollten die Möglichkeit haben, sich aufeinander einzustellen.“ (E)



Grafik 14a und 14b: Lehrkräfte - Erfahrungen mit mündlichen Klassenarbeiten

Der Vorbereitungsaufwand für die landeszentrale Erprobung der Sprechprüfung wird von den betroffenen Lehrkräften ganz überwiegend *positiv* eingeschätzt. Allerdings sind auch hier deutlich kritische Stimmen zu vernehmen, die den Aufwand als zu hoch im Vergleich zu den bisherigen Prüfungsformen empfinden und eher für eine Beibehaltung des bisherigen Prüfungsformats plädieren: „Für den Aufwand, der dafür betrieben wird, kann ich jetzt nicht unbedingt erkennen, dass es vom Wert her dann auch diesen Aufwand rechtfertigt“ (E). – „An sich eine ganz tolle Sache, aber der Organisationsaufwand ist viel zu hoch.“ (F)

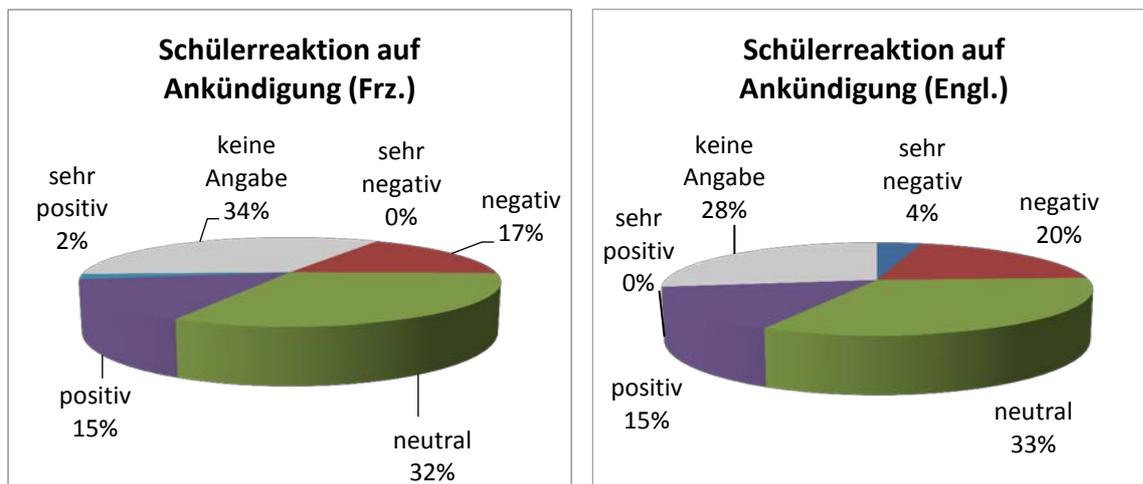


Grafik 15a und 15b: Lehrkräfte - Verhältnis von Vorbereitungsaufwand und Resultat bei den landeszentralen Sprechprüfungen

3.2.2 Reaktionen auf die Erprobung der landeszentralen Sprechprüfung

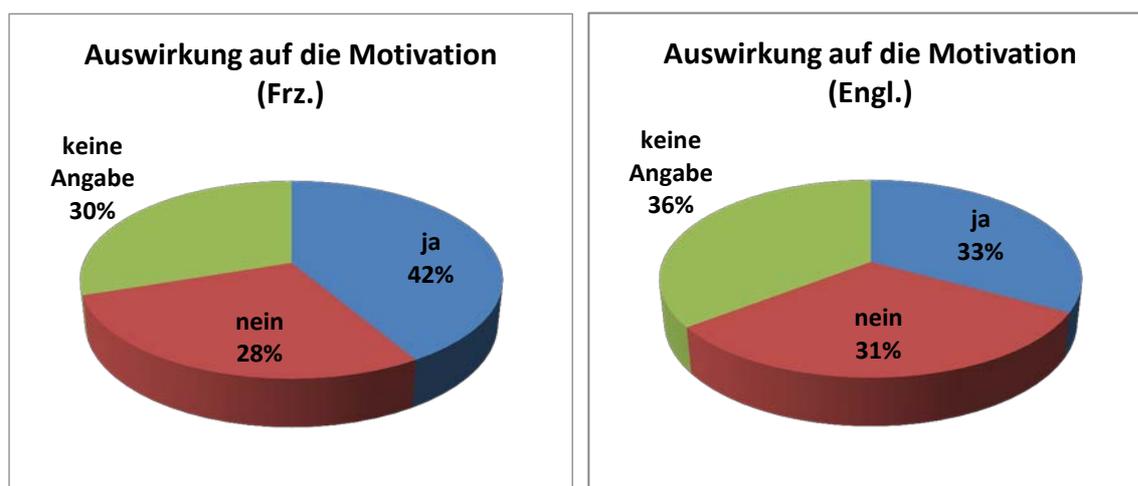
Die Angaben zu den Schülerreaktionen auf die Ankündigung der Sprechprüfung gestalten sich bei den Französischlehrkräften ausgeglichen (15% *positive*, 17% *negative* und 34% *neutrale* Reaktionen; 34% machen allerdings keine Angaben dazu).

Während der Anteil der als *neutral* und *positiv* wahrgenommenen Reaktionen bei allen Lehrkräften in etwa gleich ist (33%), ist der Anteil der als *negativ* (20%) und *sehr negativ* (4%) eingeschätzten Schülerreaktionen bei den Englischlehrkräften höher. Keine Englischlehrkraft schätzt die Reaktion als *sehr positiv* ein; der Anteil derer, die keine Angaben machen, ist auch hier mit 28% vergleichsweise hoch.



Grafik 16a und 16b: Lehrkräfte - Schülerreaktion auf die Ankündigung der landeszentralen Sprechprüfung

Positive Auswirkungen der Ankündigung auf die Motivation der Schülerinnen und Schüler werden von den Französischlehrkräften deutlicher (42%) wahrgenommen als von den Englischlehrkräften (33%), allerdings macht jeweils etwa ein Drittel der Befragten hier keine Angaben. Offenbar ist die Einschätzung der Motivation angesichts des neuen Prüfungsformats für die Lehrkräfte schwierig.



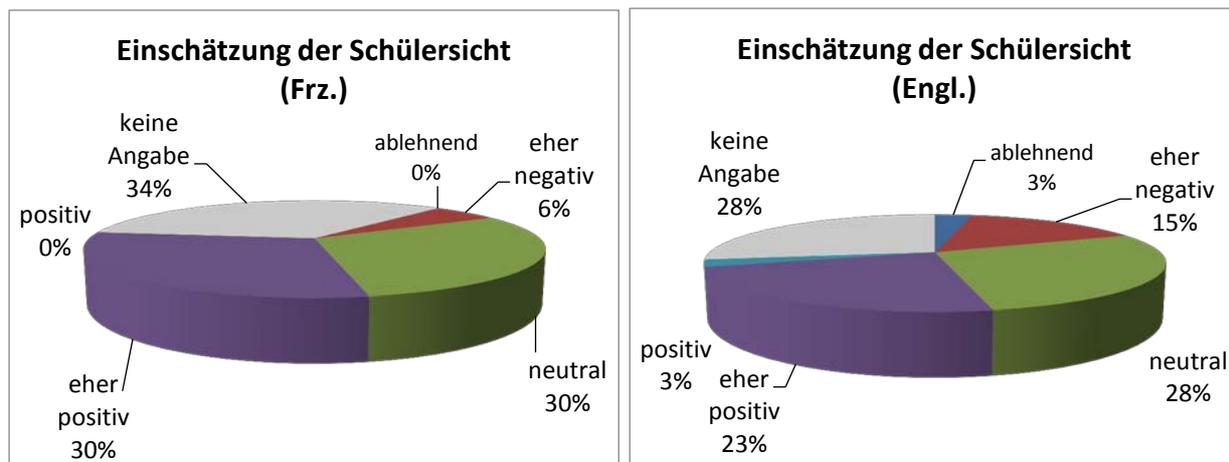
Grafik 17a und 17b: Lehrkräfte - Auswirkung auf die Motivation der Schülerinnen und Schüler

Aus der Sicht der Französischlehrkräfte, die sich hierzu geäußert haben, erlebten die Schülerinnen und Schüler die Prüfungen als *neutral* bis *positiv* (je 45,8%); nur ein geringer Anteil (8,4%) geht von einem *eher negativen* Erleben seitens der Schülerinnen und Schüler aus. Bei den Englischlehrkräften stellt sich das Prüfungserlebnis der Schülerinnen und Schüler etwas weniger positiv dar. 4,4% der auf diese Frage antwortenden Englischlehrkräfte schätzen die Sicht ihrer Schülerinnen und Schüler auf die durchgeführten Sprechprüfungen als *ablehnend* ein, 21,11% als *eher negativ*, 38,9% als *neutral*, 32,2% als *eher positiv* und 3,3% als *positiv*.

In den Freitextantworten werden einerseits eine größere Motivation zum Sprechen insgesamt und eine höhere Bereitschaft, auch mit wechselnden Partnern zu sprechen, genannt. Eine Englischlehrkraft merkt an, dass die Vorbereitungen den Schülerinnen und Schülern zeigen, „dass Reden Spaß macht“ (E); genannt wird weiterhin eine „größere Motivation zur Vorbereitung, da auch auf unbekannte Inhalte spontan reagiert werden muss.“ (E) Andererseits wird mehrfach die Unsicherheit der Schülerinnen und Schüler vor der Prüfung hervorgehoben. So führt z.B. eine Englischlehrkraft aus: „Schülerinnen und Schüler fühlten sich als

Versuchskaninchen" (E) und „waren sehr angespannt wegen möglicher schwierigen Themen" (E) Dennoch wird hier auch betont, dass die Bereitschaft zum Sprechen insgesamt durch die Notwendigkeit der Prüfungsvorbereitung erhöht wurde.

Von den Eltern haben die befragten Lehrkräfte nur ganz vereinzelt Rückmeldungen erhalten.



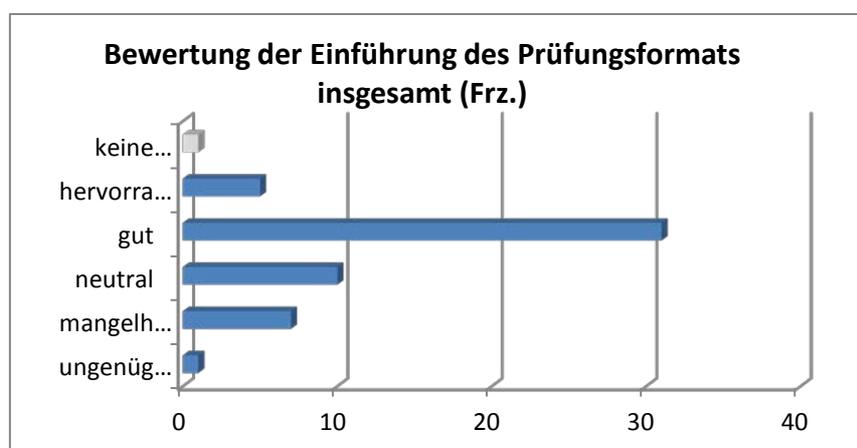
Grafik 18a und 18b: Lehrkräfte - Einschätzung der Schülersicht nach der Prüfung

3.2.3 Prüfungsdurchführung

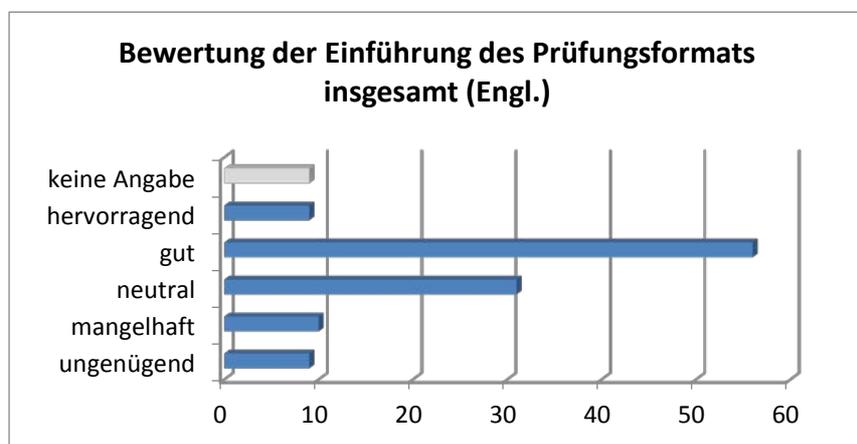
3.2.3.1 Prüfungsformat

Die Einführung der landeszentralen Sprechprüfung als neues Prüfungsformat wird in beiden Fächern begrüßt; sie werden ganz überwiegend mit *gut* bewertet. Jedoch verdienen hier die Antworten *mangelhaft* und *neutral* besondere Aufmerksamkeit. In den Freitextfragen und insbesondere auch den Interviews zeigt sich, dass eine solche Einführung von vielen Lehrkräften stark an die Forderung nach organisatorischen Veränderungen gekoppelt wird (z.B. Zeitpunkt, Taktung der einzelnen Prüfungen, längere Zeit für die Bewertung, Vorbereitungszeit für Lehrkräfte sowie Schülerinnen und Schüler), Vereinfachung der Protokollbögen; im Französischen auch hinsichtlich einzelner Themen).

Eine Englischlehrkraft fasst dies deutlich zusammen: „Ich persönlich würde es begrüßen, wenn wir eine Sprechprüfung im Abitur bekommen, da es eine Kompetenz ist, die sehr wesentlich ist und an Bedeutung gewinnt, sodass, wenn wir sie ins Abitur aufnehmen, sie die erforderliche Relevanz bekommt. [...] Jedoch mit Überarbeitung des Aufgabenformats oder einer generellen Überarbeitung. [...] Trotzdem wäre es das wert!“ (E)



Grafik 19a: Lehrkräfte - Bewertung des Prüfungsformats insgesamt (Frz.)



Grafik 19b: Lehrkräfte - Bewertung des Prüfungsformats insgesamt (Engl.)

Die Sprechprüfungen erfassen nach Einschätzung der Lehrenden weitgehend die bekannte mündliche Leistung der Schülerinnen und Schüler (F: 68%, E: 64%). Laut den Aussagen der Englischlehrkräfte ergaben sich bei den Noten im Vergleich zur letzten Zeugnisnote meist keine oder nur recht geringfügige Veränderungen (1-3 Punkte), und diese mehrheitlich zum Besseren (40,3%, im Vergleich mit 16,1% zum Schlechteren). Größere Abweichungen waren selten und dann ebenfalls meist positiv (Veränderungen um 4-6 Punkte in 4% aller Fälle, bei 1,6% besser,

Veränderungen um mehr als 6 Punkte bei 4% zum Besseren, bei 0,8% zum Schlechteren). Vergleichbare Angaben machen auch die Französischlehrkräfte.

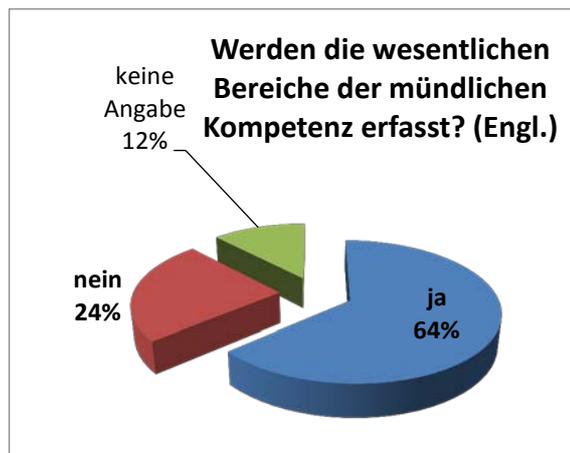
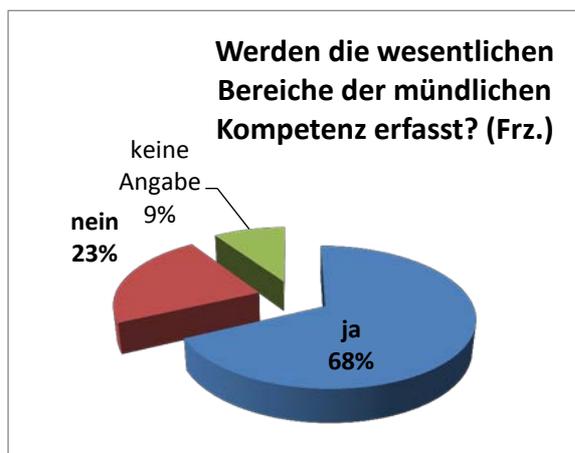
Allerdings wird mehrfach darauf hingewiesen, dass die Zeugnisnote keine gute Vergleichsgröße darstelle, da diese andere Kompetenzen abbilde (z.B. schriftliche Leistungen). Manche Lehrkräfte betonen, dass in den Sprechprüfungen Teilaspekte bewertet wurden, die schriftlich nicht bewertbar sind, und dass mündliche Prüfungen generell oft bessere Ergebnisse hervorbringen.

Allgemein wird in beiden Fächern angemerkt, dass die Prüfungen bei guter Vorbereitung recht positiv ausfielen und Schülerinnen und Schüler mit besseren kommunikativen Fähigkeiten besser abschnitten. Schlechtere Noten wurden öfter durch Nervosität oder durch bessere rezeptive Fähigkeiten der Schülerinnen und Schüler erklärt.

Hier einige Einzelstimmen:

- „Es gab Schülerinnen und Schüler, die wirklich mündlich extrem gut sind [...] Andere sind schwach, können schriftlich immer noch ein mittleres Niveau schaffen, aber mündlich brechen die dann richtig ein. Also war mein Notenspektrum viel größer. Jetzt hatte ich welche mit 14 und 15, aber auch 04 (ansonsten zwischen 07 und 12).“ (E)
- „Einige haben uns positiv überrascht.“ (F)
- „Im Großen und Ganzen waren die Leistungen vergleichbar.“ (E)
- „das Ergebnis, [war] doch sehr positiv [...], wie das eigentlich bei mündlichen Prüfungen meistens ist, dass sie etwas im Niveau über der schriftlichen Prüfung liegen. Die Schülerinnen und Schüler waren eigentlich skeptischer vor dieser landeszentralen Prüfung als es notwendig war. Das Ergebnis hat das eigentlich ganz schön bestätigt, dass das gut geklappt hat.“ (F)

Einzelne Englischlehrende monieren allerdings, dass durch die knappe Zeit eher spontanere Reaktionen abgeprüft werden, da kaum Zeit besteht, sich über die Themen ausführlich Gedanken zu machen.



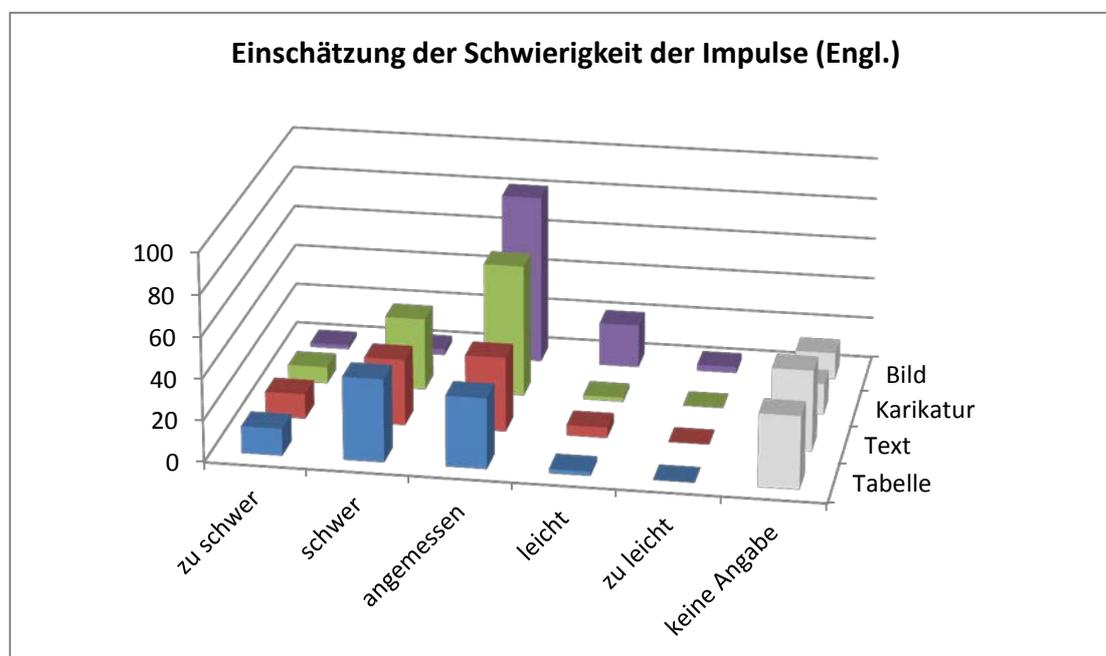
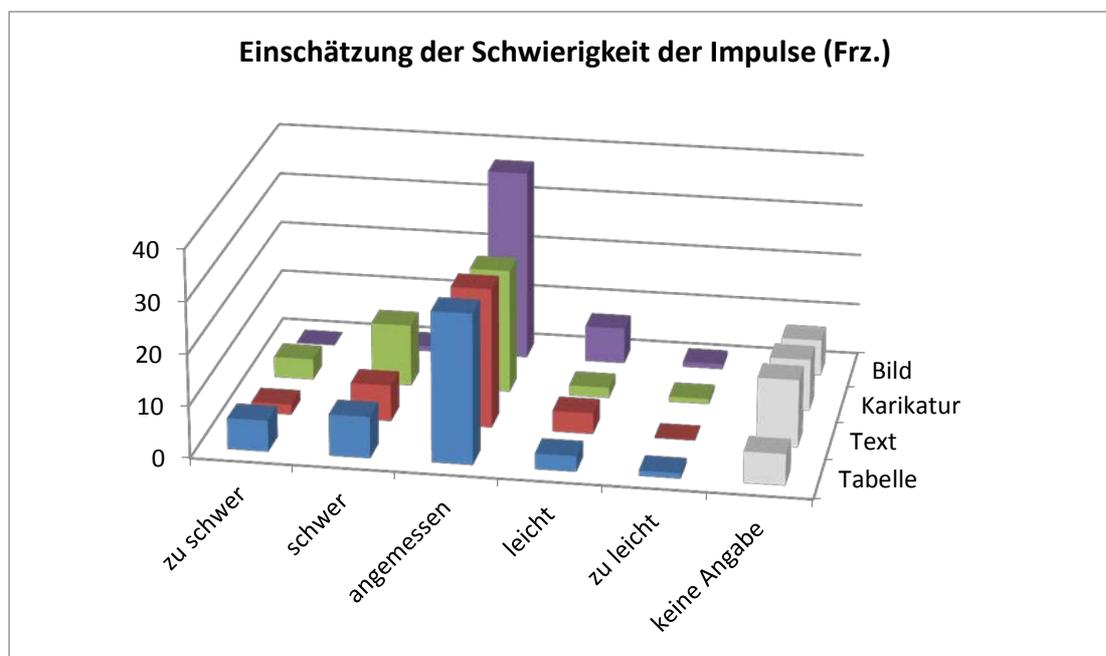
Grafik 20a und 20b: Lehrkräfte - Erfassung der wesentlichen Bereiche der mündlichen Kompetenz

3.2.3.2 Schwierigkeitsgrad

Der Schwierigkeitsgrad der einzelnen Impulse wird von den Lehrkräften differenziert gesehen. Der Bildimpuls wird im Vergleich zu den anderen Impulsen als leichter und meistens als *angemessen* empfunden, bei den Französischlehrern mit einem besonders hohen Anteil an *angemessenen* Bildern.

Hier zeigen sich Unterschiede zwischen Englisch und Französisch, wo der Impuls *Karikatur* eine noch deutlichere Tendenz in Richtung *schwer* erfährt. In den Interviews wird hier mit einem unterschiedlichen Anspruchsniveau dieser beiden Impulse argumentiert. „Das [Karikatur] hatte Niveau und das andere [Bild] war zum Beschreiben.“ (E) Eine Englischlehrkraft führt aus, dass „möglichst keine Karikaturen in einer solchen Prüfung [verwendet werden sollten] – Schülerinnen und Schüler gehen baden bei der Interpretation einer Karikatur in einer

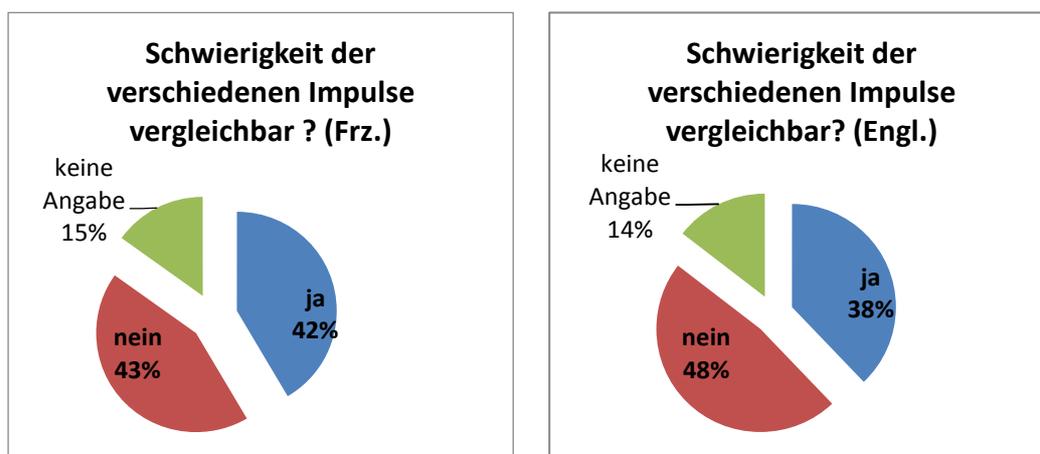
Stresssituation.“ (E) Die Englischlehrenden führen hier aus: „Texte geben zu viel Wortschatz vor; schwächere Schülerinnen und Schüler haben Schwierigkeiten, Karikaturen spontan zu erfassen“. (E) Sie kontrastieren weiterhin die „detaillierte Karikatur im Vergleich zu eindeutigen Bildern, die wenig interpretierende Beschreibungen zulassen mit ähnlichem Arbeitsauftrag.“ (E) Hier wird ergänzt, dass die Fotos teilweise zu allgemein gehalten waren. Auch wird hier angeführt, dass „der Schwierigkeitsgrad je nach Gruppe stark schwankte“ (F) und dass ein Gefälle im Bereich der Fülle an Redeanlässen bestehe, sodass in einigen Prüfungen die Schülerinnen und Schüler eine so hohe Fülle an Inhalten bearbeiten sollten, dass sie deswegen „nicht fertig wurden.“ (E) Auf der anderen Seite wird ein Mangel an Inhalt kritisiert, der dazu geführt habe, „dass sie nicht wussten, was sie noch erzählen sollten.“ (E)



Grafik 21a und 21b: Lehrkräfte - Einschätzung der Schwierigkeit der Impulse Bild - Karikatur - Text - Tabelle

Auch bei der Vergleichbarkeit des Schwierigkeitsgrads der Impulse gibt es unterschiedliche Wahrnehmungen, was besonders in den Interviews deutlich wird. Die Vergleichbarkeit leidet nach Ansicht der Lehrkräfte unter den oben beschriebenen Sachverhalten. Eine klare Position in dieser Frage kann über alle befragten Lehrkräfte hinweg nicht festgestellt werden. Dies wird durch folgende Einzelstimmen deutlich:

- „In Französisch waren die Aufgaben vom Schwierigkeitsgrad durchaus unterschiedlich, zum Teil Karikaturen mit Text oder ohne Text, zum Teil Statistiken, wo man einfach mehr Zeit braucht, um sich einzulesen.“ (F)
- „Karikaturen sind grundsätzlich schwierig, weil die Zeit nicht da ist.“ (E)
- „Bei Diagrammen und Statistiken sollten nicht noch andere Impulse gleichzeitig gegeben werden.“ (E)
- „Also in Englisch muss ich sagen, die Prüfungen hatten prinzipiell alle einen ordentlichen Anspruch.“ (E)
- „Die Prüfungen waren sehr unterschiedlich angelegt, ... ich hätte mir hier ein größeres Gleichgewicht gewünscht.“ (E)



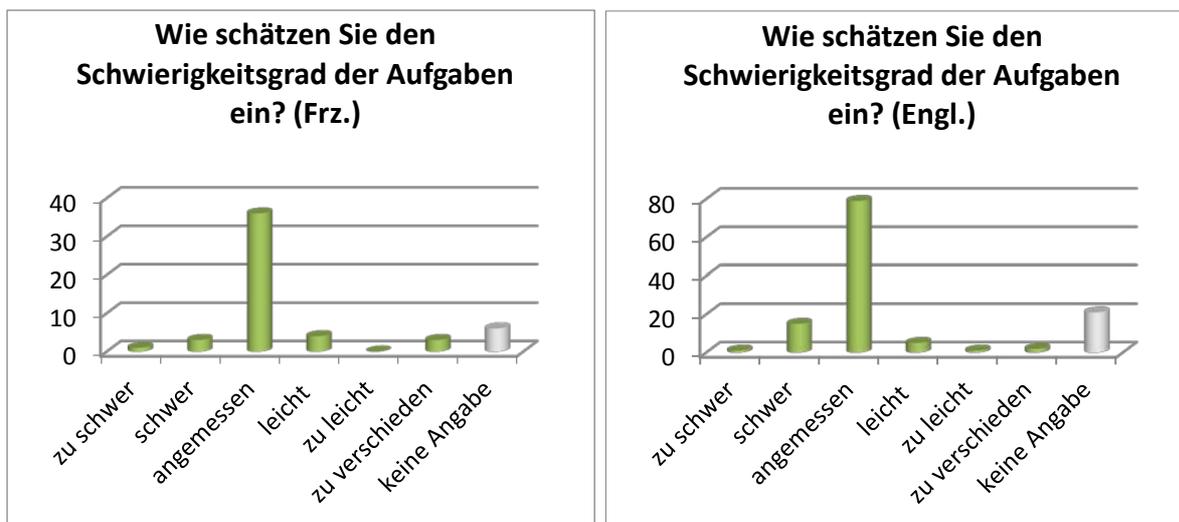
Grafik 22a und 22b: Lehrkräfte - Vergleichbarkeit der verschiedenen Impulse

Insgesamt wird der Schwierigkeitsgrad der Aufgaben von den Lehrkräften beider Fächer ganz überwiegend als *angemessen* empfunden, wie Grafik 23 zeigt. In den Interviews werden ergänzend verschiedene Aspekte thematisiert, etwa,

- dass „[d]ie Aufgaben [...] zuvor sehr anspruchsvoll [waren], während in der Prüfung meine 3er Gruppe Probleme hatte über das Thema 12 Minuten lang zu sprechen.“ (F)
- „[...], dass sie [die Schülerinnen und Schüler] die Themen aus der Prüfung einfacher und leichter zugänglich empfanden als in der Vorbereitung. Die Themen in der Prüfung waren eher allgemein gehalten, wohingegen in der Übung diese sehr spezifisch waren. Das wurde als positiv empfunden und von den Schülerinnen und Schülern gutgeheißen.“ (E)

Bei intensiver Vorbereitung auf das Prüfungsformat im Hinblick sowohl auf Inhalte als auch auf Abläufe wurde der Schwierigkeitsgrad über beide Fächer hinweg als *angemessen* bewertet. Dennoch gibt es Stimmen, die monieren: „Einzelne Aufgaben lagen vom Schwierigkeitsgrad her für die Schülerinnen und Schüler doch deutlich über den anderen.“ (F) Teilweise wird dies mit der Frage der Vorbereitungszeit in Zusammenhang gebracht. So erklärt eine Englischlehrkraft: „Einfach sehr viel an Input! Man braucht ja auch eine gewisse Zeit, egal wie gut man ist, diese Bilder und Grafiken umzusetzen; hier sollte man sich die Frage stellen: Was ist unser Ziel? Im Prinzip geht es um die sprachliche Leistung und nicht um die Fähigkeit innerhalb weniger Sekunden zu interpretieren.“ (E) Aufgrund mangelnder Einlesezeit würden auch viele Impulse nur auf dem Niveau des Beschreibens umgesetzt oder weitere Redeanlässe bzw. Themenfelder der vorliegenden diskursiven Thematik unbeachtet bleiben.

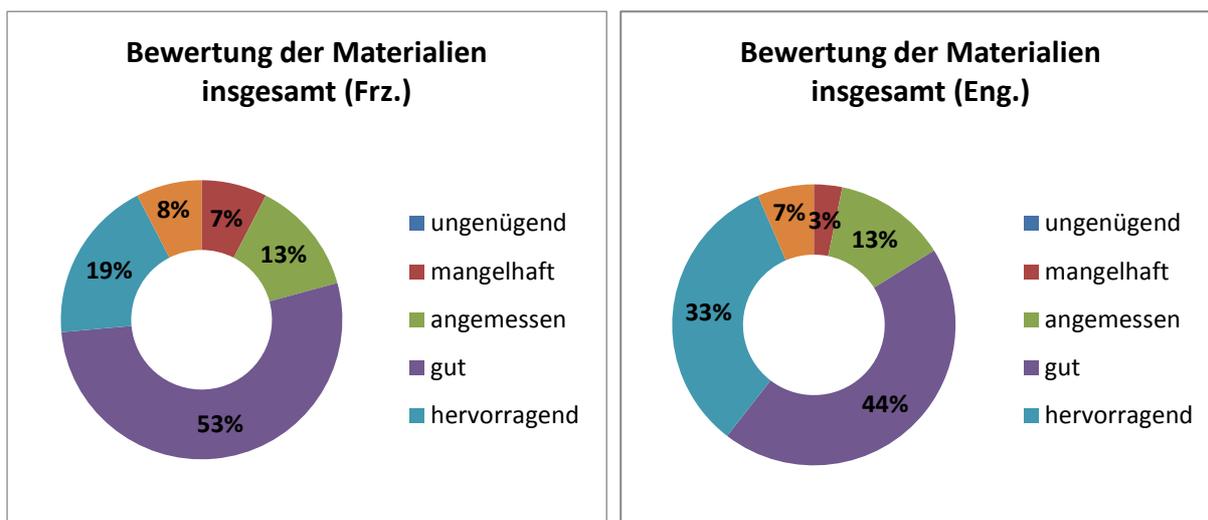
Gerade im Fach Französisch mit seinen exakten thematischen Vorgaben im gymnasialen Oberstufenlehrplan fällt dies auf. Obwohl im Fach Englisch diese Problematik in den Antworten weniger zu erkennen ist, wird auch hier angemerkt: „Manche Themen kommen ohne viel Vorwissen und Themenvokabular aus, andere verlangen den Schülerinnen und Schülern Einiges ab.“ (E)



Grafik 23a und 23b: Lehrkräfte - Einschätzung des Schwierigkeitsgrades der Aufgaben

3.2.3.3 Materialien und Inhalte

Die zusammenfassende Bewertung der Materialien ist insgesamt positiv; 72% der Französischlehrkräfte und 77% der Englischlehrkräfte bewerten sie als *gut* oder *hervorragend*.



Grafik 24a und 24b: Lehrkräfte - Bewertung der Prüfungsmaterialien und -inhalte insgesamt

Dies wird auch in den Interviews deutlich, in denen sich aber auch punktuelle Kritik zeigt, z.B. an der Formulierung der Aufgabenstellung. Außerdem wünschen sich einige Lehrende eine frühere Einsicht in die Materialien und das Prüfungsskript:

- „Es wäre wünschenswert gewesen, „wenn wir uns hätten vorbereiten können oder wenigstens mal eine Stunde vorher über die Materialien schauen können, damit wir wussten, was auf uns zukommt.“ (E)
- „Ich gehe blind in eine Prüfung rein! Ich habe keine Möglichkeit mir irgendwo richtig vorher Gedanken zu machen“ (F)

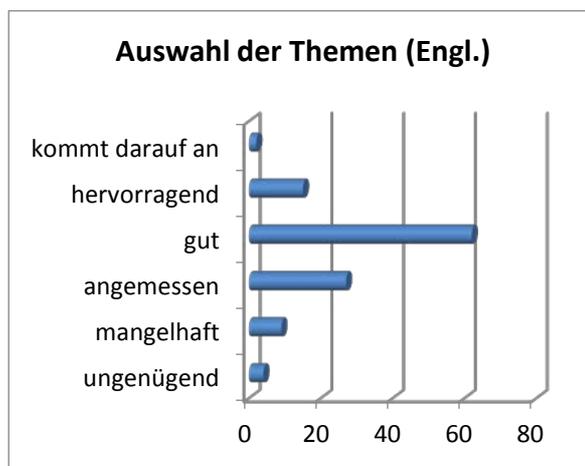
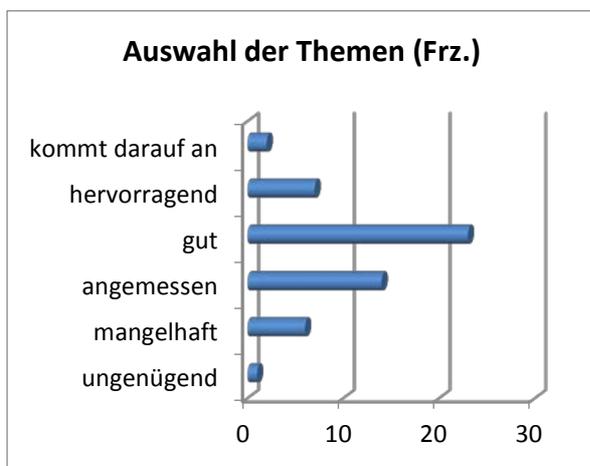
- „Und da wäre es gleich hilfreich, wenn die Lehrer die Materialien auch früher bekommen, gerade wenn ein Gespräch, wie bei den Cartoons, zum Erliegen kommt [...] Ich weiß ja dann, das könnte vielleicht problematisch werden, wie bringe ich das Gespräch dann wieder in Gang.“ (E)

Ein relativ hoher Anteil von Lehrenden (26% bei den Französisch-, 30% bei den Englischlehrkräften) stuft die Aufgabenstellung als nicht auf Anhieb erkennbar ein. Die Englischlehrkräfte führen hierzu aus, dass die Aufgaben inhaltlich zu komplex strukturiert seien und zu viele Teilaspekte und Arbeitsanweisungen enthielten und dass viele Details und Operatoren in der kurzen zur Verfügung stehenden Vorbereitungszeit erfasst werden müssten. Eine Lehrkraft formuliert anschaulich: „Much too much, zu viel Text, schwer spontan zu reagieren in so kurzer Zeit.“ (E) Vereinzelt wird angemerkt, dass Bilder und andere Aufgaben verschieden interpretierbar seien (s.o.) und dass Nervosität bei den Schülerinnen und Schülern zu Übersehen oder Fehlinterpretation führe. Insgesamt seien die Aufgaben aber „gut machbar: beschreiben, vergleichen, kontrastieren; ergänzbar durch eigene Erfahrungen.“ (E)



Grafik 25a und 25b: Lehrkräfte - Wurden die Aufgaben auf Anhieb klar?

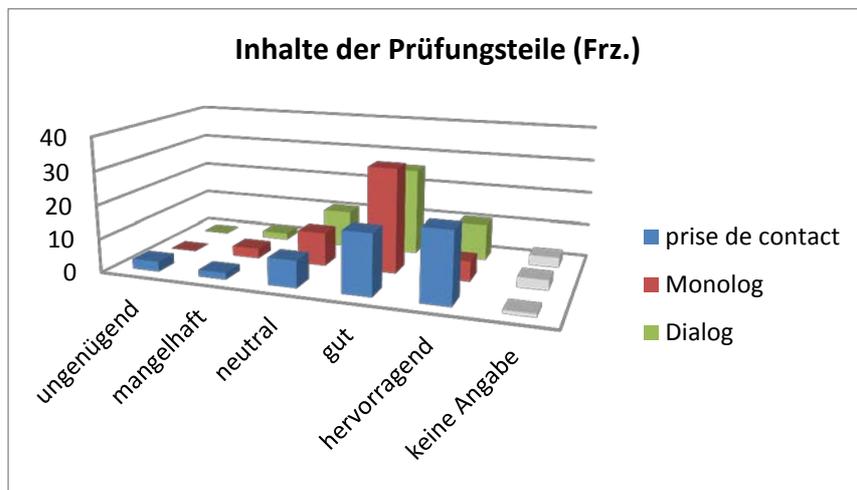
Die Themenauswahl wird von Lehrenden beider Fächer ganz überwiegend *angemessen* bis *hervorragend* eingestuft; negative Meinungen sind im Vergleich dazu eher selten. Wenn diese geäußert werden, wird als Begründung etwa angeführt, dass die Themen aus den Vorbereitungsmaterialien zum Teil nicht kongruent mit den Inhalten der landesweiten Sprechprüfung bzw. nicht im Lehrplan enthalten waren. Auch in den Interviews wird diese Frage angesprochen, so z.B.: (...) [Themen] waren nicht vollständig deckungsgleich mit den Materialien zur Übung, z.B.: *sécurité routière*.“ (F) Einige Lehrende merken an, dass manche Aufgaben nicht der Lebenswelt der Schülerinnen und Schüler entsprächen, wobei andere Lehrende demgegenüber gerade die Themenauswahl als angewandt und lebensnah loben.



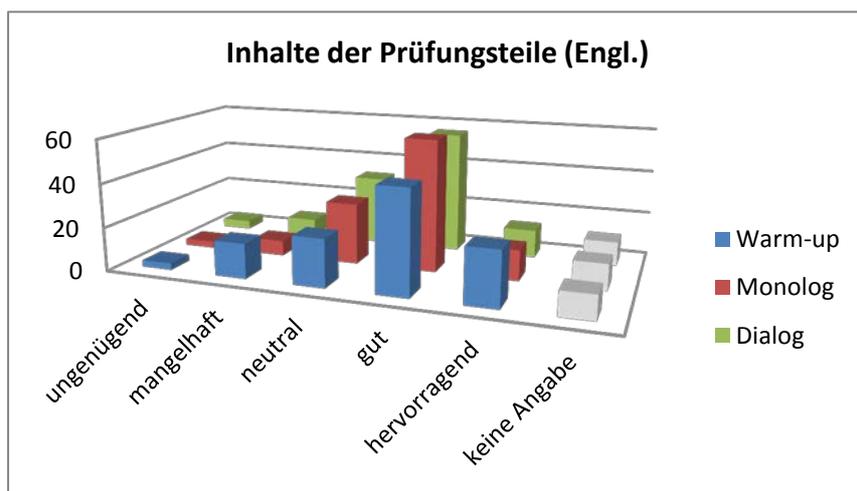
Grafik 26a und 26b: Lehrkräfte - Themenauswahl

3.2.3.4 Prüfungsablauf

Die einzelnen Prüfungsteile Warm-up/Prise de contact, Monolog und Dialog werden von den Lehrkräften fächerübergreifend ganz überwiegend als *gut* bis *hervorragend* bewertet. Allerdings findet sich unter den Englischlehrkräften mit 13% ein relativ hoher Anteil, der den Warm-up-Teil als *mangelhaft* einschätzt.



Grafik 27a: Lehrkräfte - Prüfungsteile Französisch im Vergleich



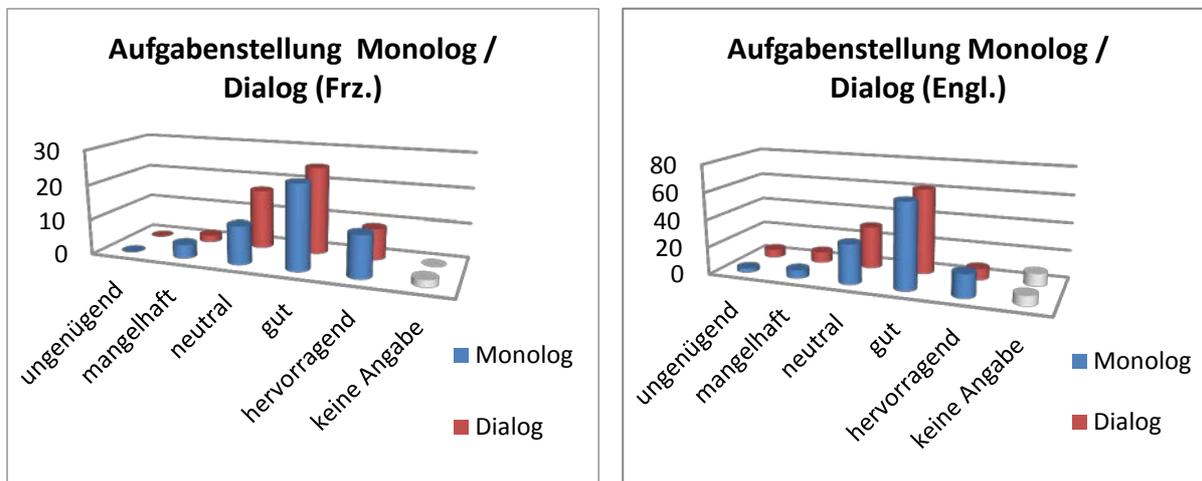
Grafik 27b: Lehrkräfte - Prüfungsteile Englisch im Vergleich

Die Aufgabenstellungen selbst (sowohl monologischer als auch dialogischer Teil) erhalten von Englisch- und Französischlehrenden *gute* bis *hervorragende* Bewertungen (63,7% der Englischlehrkräfte beim monologischen, 56,5% beim dialogischen Teil; Französischlehrkräfte 68% bzw. 64,2%). Beim monologischen Teil wird positiv bewertet, dass die Aufgabenstellung gut und schnell zu erfassen und zu bewältigen sei. Eine Lehrkraft erläutert: „Die Schüler wussten in weiten Teilen, welche Anforderungen auf sie zukommen und waren vertraut mit den Operatoren ‚compare and contrast.‘“ (E) Die „Beschreibung des visuellen Materials gibt Sicherheit“ (E); weiterführende Impulse werden allerdings als „nicht immer ergiebig“ (F) eingeschätzt, insbesondere bei Bildern.

Der dialogische Teil wird von einigen Englischlehrenden als komplex gesehen. Hier merken Lehrkräfte mehrfach an, dass die Aufgabenstellungen im Kontext der zur Verfügung gestellten Zeit sehr umfangreich und anspruchsvoll seien: Das „task fulfillment“ sei hier „in der Kürze der Zeit nicht immer leistbar;“ (E) die Schülerinnen und Schüler seien aber „meistens in der Lage, zu

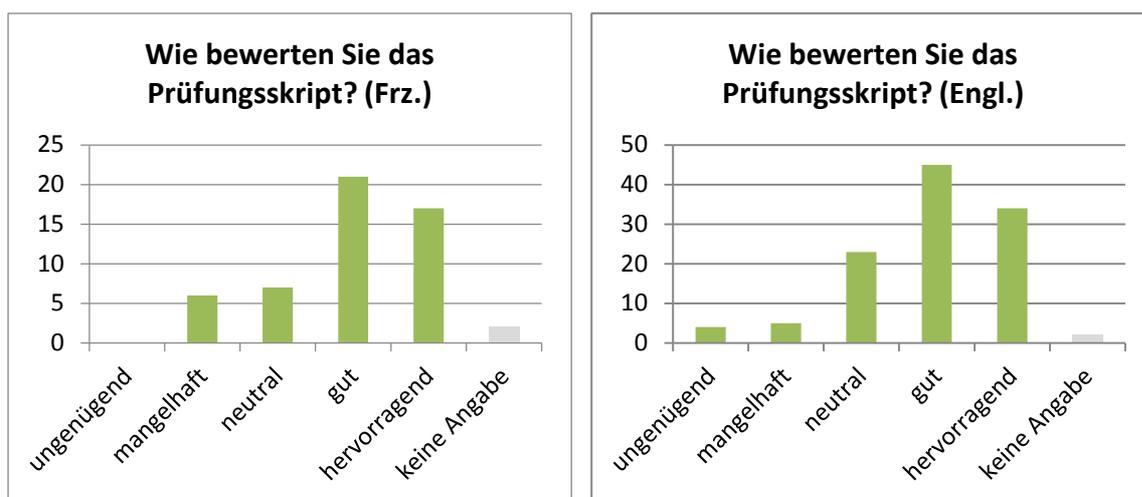
vorgegebenen Aufgaben Ideen zu entwickeln,” (E) wobei „öfters zusätzliche Impulse notwendig” (E) seien.

Ein weiterer Kritikpunkt ist die Interaktionskomponente. In Einzelmeinungen und Interviews weisen Lehrende auf die Interaktionsproblematik gerade im monologischen Teil hin: „Der Teil der Interaktion ist nicht immer realisierbar, gerade im Monolog; gerade hier in dem Bereich Interaktion und Vortrag haben viele Schülerinnen und Schüler weniger Punkte bekommen, weil es eben wirklich schwer zu verwirklichen ist. Wie soll man denn viele produktive Gesprächsimpulse liefern, wenn das eigentlich ein Monolog sein soll – Das passt für mich nicht!“ (E)



Grafik 28a und 28b: Lehrkräfte - Aufgabenstellung Monolog/ Dialog im Vergleich

Allgemein wird die Existenz eines Prüfungsskripts, das die Vergleichbarkeit der Noten unterstützt, sehr begrüßt. Jedoch werden weniger starre Vorgaben gefordert, so dass kleine Freiräume erhalten bleiben, die es ermöglichen, situationsgerecht auf die Prüflinge zu reagieren. Dies sollte nach Ansicht einiger Lehrkräfte insbesondere auf Ebene der verbindlichen Zusatzimpulse geschehen. Obwohl diese einerseits als „ideal” (E) und „klasse”(E) bezeichnet werden, sollten andererseits Regelungen wie „höchstens einen oder zwei [Zusatzimpulse]“ (E) natürlicher gestaltet werden und eher als „Handreichung mit möglichen Beispielen“ (E) bzw. „mehr als Möglichkeit als ein Muss“ (E) formuliert werden. Denn: „Das kommt eigentlich spontan, da muss ich nicht auf ein Blatt schauen. Aber es stört mich nicht, wenn es da steht.“ (E) Auch wird vorgeschlagen, die Prüfungsskripte früher zu veröffentlichen, damit sich die Lehrkräfte besser vorbereiten können.



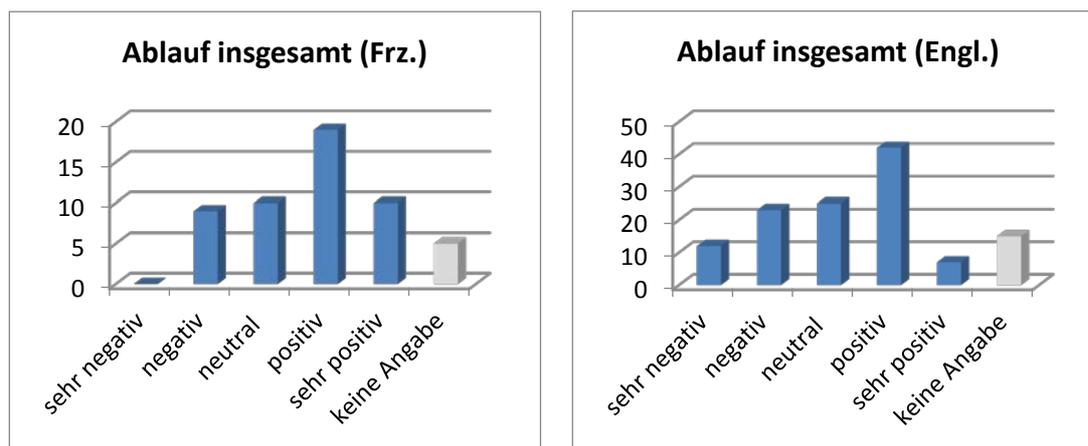
Grafik 29a und 29b: Lehrkräfte - Bewertung des Prüfungsskripts

Der Ablauf des Probedurchgangs der landeszentralen Sprechprüfung wird insgesamt überwiegend positiv eingeschätzt, bei den Englischlehrkräften allerdings etwas weniger deutlich als bei den Französischlehrenden. Bei der Frage nach Problemen, die die Lehrkräfte hinsichtlich der Prüfung sehen, werden zahlreiche Einzelaspekte genannt: Am häufigsten wird die zu enge zeitliche Taktung der Prüfungen kritisiert, die damit als Stresssituation für Prüfende wie Prüflinge empfunden wird:

- „Es gibt zu viele Bereiche, auf die man achten muss; man kann mit dem Zweitprüfer dies kaum besprechen – Man kann nicht Protokoll bearbeiten, Beispiele aufschreiben und Punkte vergeben... Das ist unmöglich!“ (E)
- „Unglaubliche Papierfülle, zu wenig Zeit für Bewertung, Verzug, Aufwärmphase 'leidet', man beachte, dass Schülerinnen und Schüler je ca. 1 Min zum Betreten und Verlassen des Prüfungsraumes brauchen, Kollegen, die nicht 100% fit sind, sind überfordert, extremer Stress durch enge Taktung.“ (E)
- „Viel Organisationsaufwand, Zeit fürs Protokollieren war viel zu kurz. Es war anstrengender als die normale mündliche Kursarbeit.“ (E)
- „[...] von erster bis zur letzten Prüfung, diesen Kollegen waren keine geraden Gedanken mehr möglich.“ (F)

Aber auch gegenteilige Meinungen sind vertreten: „[D]as ging. Ich habe mich nicht gestresst gefühlt.“ (F)

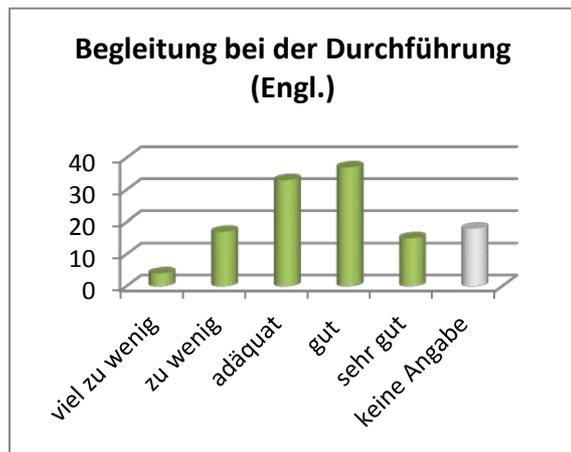
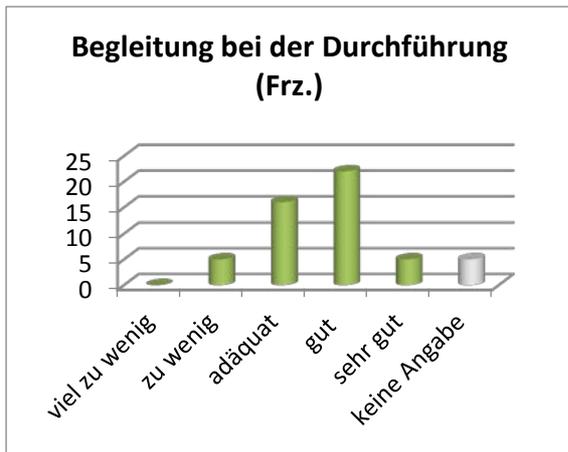
Abhilfe könne neben der bereits erwähnten längeren Vorbereitungszeit (vgl. Kap. 3.2.3.2)⁴ z.B. auch die frühere Bekanntgabe der jeweiligen Prüfungspartner schaffen: „Die Schülerinnen und Schüler könnten ruhig früher erfahren, wer ihr Partner ist und auch wann die Prüfung ist; das entspannt wesentlich, von der Logistik und es entspannt vom Kopf her.“ (F)



Grafik 30a und 30b: Lehrkräfte - Ablauf insgesamt

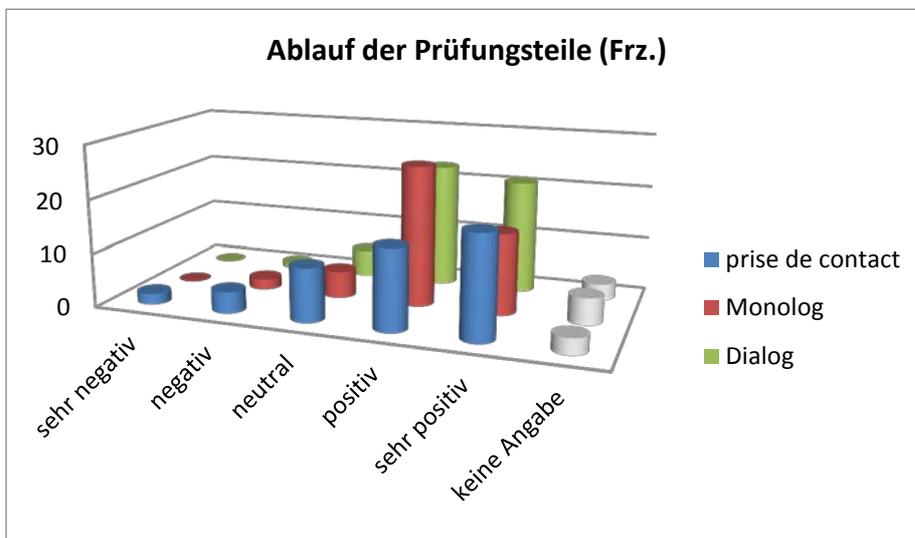
Auch hinsichtlich der Begleitung der Prüfungsdurchführung ist eine insgesamt positive Tendenz bei den befragten Lehrkräften festzustellen, die meisten fühlen sich *adäquat* bis *gut* begleitet.

⁴ Ggf. auch mit Hilfsmitteln: Angesichts der zwei für die Aufsicht im Aufenthaltsraum anwesenden Personen könne die Aufsicht auch bei der Vorbereitung von Schülerinnen und Schülern gewährleistet werden, so eine Lehrkraft: „Ich würde auch erlauben, dass sie Hilfsmittel in der Vorbereitung benutzen dürfen, weil das der Realität entspricht, heute haben alle ein Handy, können nachschauen.“ (F)

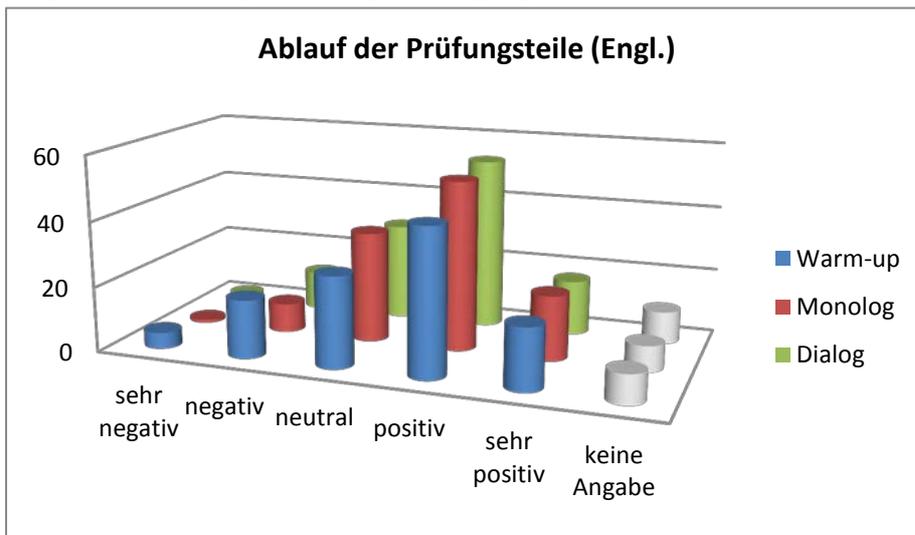


Grafik 31a und 31b: Lehrkräfte - Bewertung der Begleitung bei der Durchführung

Der Ablauf der einzelnen Prüfungsphasen Warm-up/ Prise de contact, monologischer und dialogischer Teil sowie Notenfindungsphase wird von den Lehrkräften unterschiedlich eingeschätzt:



Grafik 32a: Lehrkräfte - Ablauf der Prüfungsteile Französisch



Grafik 32b: Lehrkräfte - Ablauf der Prüfungsteile Englisch

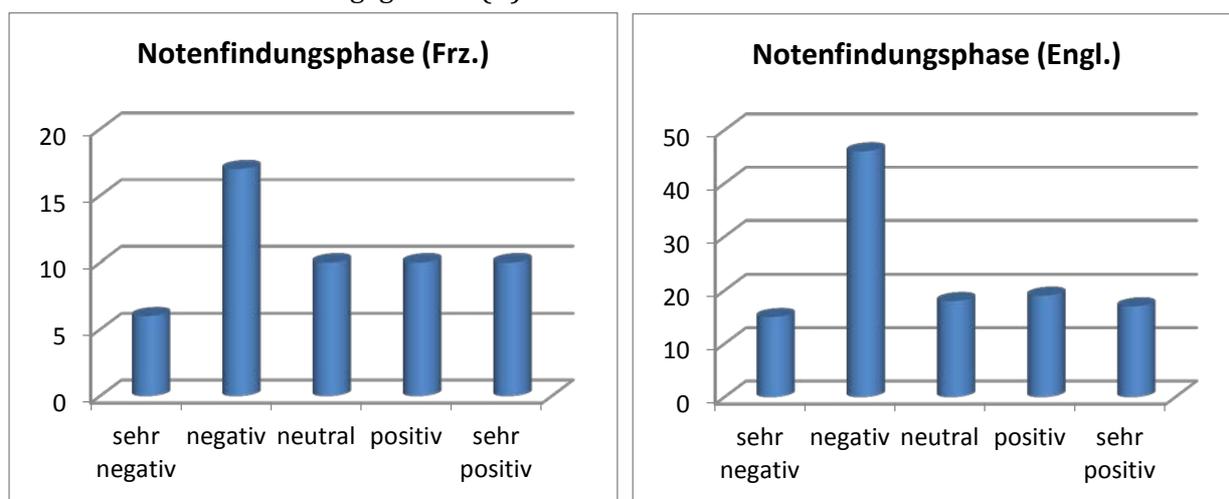
Während die eigentlichen Prüfungsteile überwiegend positive Resonanz haben, fällt die negative Einschätzung der Notenfindungsphase auf. Vor allem die kurze Zeitspanne von etwa 5 bis max. 10 Minuten für die Notenfindung wird als negativ und zusätzlicher Stressfaktor für die Prüfer angesehen, wie diese ausführen:

- „Man hatte 10 Minuten, um alle Kriterien zu besprechen. Früher war das einfacher. Jetzt allen Kriterien nachzugehen und zu entscheiden, ist es 3 Punkte oder 4 Punkte? Also jetzt ging das ganz glatt, weil wir das mit dem Kollegen natürlich auch schnell einigen konnten. Aber wenn man sich nicht einigen kann, reichen die 10 Minuten auf keinen Fall. [...] Da standen wir unter Zeitdruck.“ (F)
- „Es gibt zu viele Bereiche, auf die man achten muss; man kann mit dem Zweitprüfer dies kaum besprechen – Man kann nicht Protokoll bearbeiten, Beispiele aufschreiben und Punkte vergeben... Das ist unmöglich!“ (F)
- „Doppelt auszufüllen, ... Also was ich gemacht hatte: Es gab ja das Raster und das Protokollblatt alles in A4, stellen Sie sich vor wie dieser Tisch aussah bei einer 3er Prüfung! Material für Schülerinnen und Schüler, Impulse, wenn die Schülerinnen und Schüler nicht weiterkommen. Am Ende benötigten wir einen weiteren Tisch und alles lag voll! Und dann beobachten Sie die beiden Schülerinnen und Schüler und füllen aus. Wahnsinn! Da sich zurechtzufinden ... Im Abitur hat man normal zwei Prüfer und einen Protokollanten, hier waren wir: die Prüfer, Schriftführer, Impulsgeber und haben am Ende über die Note entschieden. Sehr komplex in dieser Weise mit so vielen Blättern.“ (F)

Der Zeitmangel ist nach Angaben der Lehrkräfte bei eingespielten Bewertungsteams leichter auszugleichen als bei neu gebildeten Lehrerteams, was in einem der Interviews verdeutlicht wird: „Da wir alle sehr eingespielt sind, konnten wir die Sachen gegeneinander abwägen. Das funktioniert relativ gut.“ (F)

Die Lehrkräfte fürchten im Hinblick auf die knappe Taktung, die vollen Prüfungstage und die komplexen Bewertungsbögen in diesem Kontext auch, dass unter den gegebenen Bedingungen die Bewertung nicht vollkommen fair sein kann:

- „Dieser 30 Minuten-Rhythmus, das war ein strammes Programm. Wir hätten mehr Pausen gebraucht“ (F)
- „Wenn die Bewertung so knapp bleibt, dann hat man keine Zeit, als Prüfer sich das nächste Thema der folgenden Prüfung anzusehen. Das war wirklich schlecht. Man ist ja auch nur ein Mensch und beim letzten Prüfungspaar, da ist es sehr schwer sich noch zu konzentrieren, um adäquat zu prüfen.“ (E)
- Das ist einfach zu viel! Da denk' ich mir, wenn dieser Schüler bei mir die letzte Prüfung hat, kann ich da gewährleisten, dass dieser eine genauso klare Beurteilung erfährt als der erste? Das ist nicht gegeben.“ (E)



Grafik 33a und 33b: Lehrkräfte - Bewertung der Notenfindungsphase

3.2.3.5 Bewertung

Mit dem Bewertungsraster sind die Lehrkräfte überwiegend gut vertraut. Es wird mehrheitlich *neutral* bis *eher geeignet* eingeschätzt. Auffallend ist hier der Anteil von deutlichen Kritikern unter den Englischlehrkräften; hier empfinden es 46% als *eher geeignet* oder *geeignet*, aber immerhin 27% als *ungeeignet* oder *eher ungeeignet*. Die Hälfte (50,9%) der Französischlehrenden hingegen bewertet das Raster als *eher geeignet* und lediglich rund 15 % geben *eher ungeeignet* an.

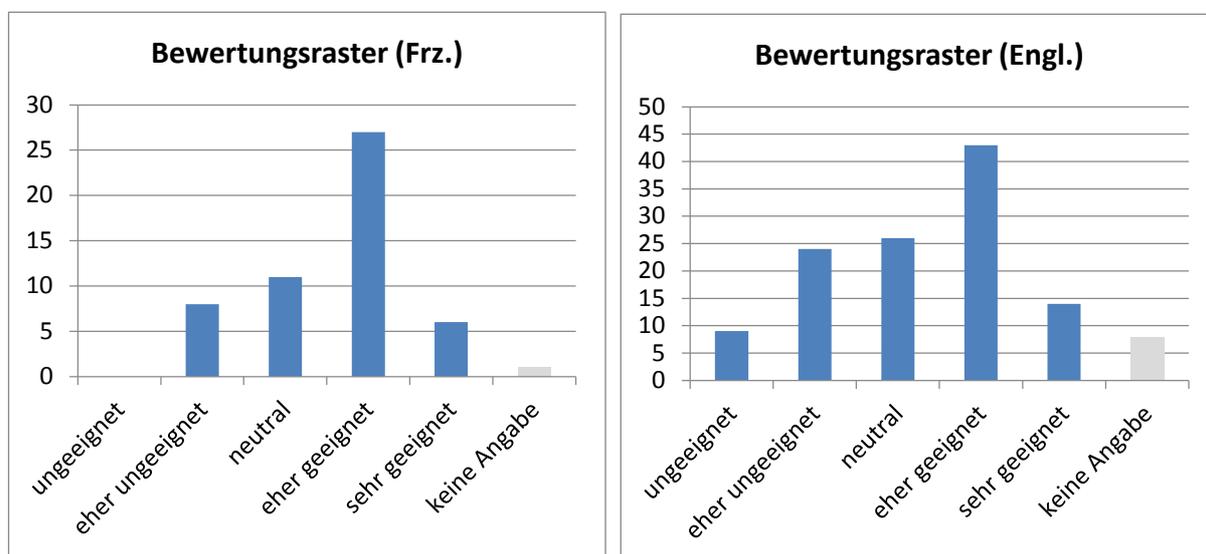
Negative Anmerkungen aus den Interviews thematisieren die nicht immer klar genug abgegrenzten Bewertungskriterien – hier wird insbesondere die mangelnde Trennschärfe zwischen Bandbreite (*Range*) und Korrektheit (*Accuracy*) genannt und auf die Schwierigkeit der Gewichtung hingewiesen. Demgegenüber heben positive Stimmen die transparente Bewertung durch das vorgegebene Schema hervor, das die Sprechniveaus gut und übersichtlich abbilde.

Mehrfach wird die enorme Papierflut angesprochen, die schnell unübersichtlich werde. Hinzu komme die Erstellung des Protokolls, wodurch die Prüfer in der Prüfung mehrere Rollen einnehmen müssen: „Prüfer, Schriftführer, Impulsgeber und haben am Ende über die Note zu entscheiden“ (F; vgl. auch oben die Aussagen zur Notenfindungsphase). Diese Problematik wird nach Angaben der Lehrkräfte in Dreierprüfungen noch deutlicher. Vielfach wird angeregt, einen zusätzlichen Protokollanten einzusetzen oder die Arbeitsfelder zwischen den Prüfern aufzuteilen. Dies würde sich nach Ansicht einiger Befragter auch positiv auf den rechtlichen Rahmen auswirken: „Man muss die Bögen halt danach noch so fertig machen, dass die einer Prüfung standhalten. Das Grobe habe wir in diesen 10 Minuten gemacht, aber trotzdem saßen wir nach den Prüfungen noch eineinhalb Stunden da und haben das Ganze dann in Form gebracht.“ (F) Eine andere Anmerkung bezieht sich auf die Schülerrückmeldung: „Dann sollte man überlegen, wenn man es verbindlich macht, welche Form der Rückmeldung stattfindet. In jeder anderen Form der Klassenarbeit ist es auch nicht nur die Note, sondern es gibt zumindest die Möglichkeit die Punkte selber nochmal nachzuzählen, was hier nicht der Fall war. Die Schülerinnen und Schüler bekamen ein Kärtchen in die Hand und da stand dann 11 oder 12 Punkte drauf.“ (F)

Auch wird hier häufig der Zeitdruck/Zeitmangel beim Ausfüllen des Bewertungsrasters während der Prüfung angeführt:

- „Insofern wäre das für uns natürlich auch klasse, wenn wir aufzeichnen könnten. Das wäre für uns schon ein Vorteil. [...] Dann ist ja zu erwarten, dass da mal einer nachfragt. Denn wer will nachher noch anhand eines solchen Protokolls sagen, es ist 08 oder 09.“ (F)
- „Das Raster ist verwendbar, das ist gut. Ich fand auch gut, dass wir nicht mehr die unterschiedliche Gewichtung drin hatten mit den Faktoren. Es war jetzt wirklich sehr eindeutig zu bewerten, da sind wir gut mit klargekommen.“ (F)
- „[D]as reicht von der Zeit her nicht. Ich muss mich eigentlich auf die Prüfung konzentrieren. Ich tauche dann immer ab, schreibe etwas hin und verpasse letztendlich auch andere Dinge. Es müsste da wiederum eine Möglichkeit bestehen, dass man einfach manche Aspekte ankreuzen kann.“ (F)

Einige Lehrkräfte sprechen die unterschiedlichen Raster für Englisch und Französisch an und regen eine formale Vereinheitlichung an.



Grafik 34a und 34b: Lehrkräfte - Bewertung des Bewertungsrasters

Auch wenn das für die Erprobung der landeszentralen Sprechprüfungen eingesetzte Losverfahren insgesamt überwiegend positiv bewertet wird, polarisiert es die Lehrkräfte deutlich. Von den Englischlehrkräften bewerten es 43,6% als *gut* oder *hervorragend*, aber auch 25,8% als *ungenügend* oder *mangelhaft*. Unter den Französischlehrkräften geben 50,9% *gut* oder *hervorragend*, aber auch 22,6% *ungenügend* oder *mangelhaft* an.

Als Vorteile werden hier genannt, dass das Zufallsprinzip als gerechter empfunden und soziale Ausgrenzung vermieden werden könne, diese Konstellation mehr Dynamik, Flexibilität und Spontaneität der Prüflinge bringe und eine authentischere Kommunikationssituation schaffe.

Negative Aspekte des Losverfahrens sind aus Lehrkräftesicht die erhöhte Verunsicherung und der zusätzliche Stress für die Prüflinge, ebenso deren gegenseitige Abhängigkeit, die sich insbesondere bei unterschiedlichen Sprachniveaus der Schülerinnen und Schüler und mangelnder Empathie negativ auswirken und in Einzelfällen zur Benachteiligung einzelner Prüflinge führen könne. Viele Lehrende sprechen in der Befragung sowohl positive als auch negative Aspekte an und zeigen damit eine sehr differenzierte Wahrnehmung:

- „Positiv: gegenüber Neigungspaaren: heterogene Paare, realitätsnah; Nachteile: schürt Aufregung unter Schülerinnen und Schülern bzw. Angst, das Gespräch mit Schülerinnen und Schülern nicht aufrechterhalten zu können.“ (E)
- „Grundsätzlich gut, dass Schülerinnen und Schüler sich auf beliebige Gesprächspartner einstellen müssen (authentisch), Losverfahren birgt Ungerechtigkeiten, z.B. schwieriger mit schwachem Partner zu diskutieren.“ (E)

In den Interviews berichten Lehrkräfte von möglichen negativen Konstellationen und mangelnden Korrekturmöglichkeiten für solche „unglückliche Paare“. Dabei werden mehrmals die negativen Auswirkungen auf die Notenfindung speziell bei sehr unterschiedlichen Prüfungspaaren angesprochen: „Wenn eine schwache Schülerin/ein schwacher Schüler mit einer guten Schülerin/einem guten Schüler zusammen kommt, dann kann die gute Schülerin/der gute Schüler dem schwachen Impulse geben, sodass die schwache/der schwache etwas besser abschneidet als sonst und die starke/der starke nicht bis zu seinem Leistungsmaximum kommt. Das könnte dazu führen, dass eine Note nicht ganz optimal ist.“ (F) Allerdings: „Katastrophen sind keine eingetreten.“ (F)

Es wird von einigen Lehrkräften angeregt, das Losverfahren früher als im Probedurchlauf durchzuführen und die Prüfungsgruppen bekanntzugeben, damit den Schülerinnen und Schülern die Möglichkeit des Austauschs gegeben wird und damit verbundene Ängste abgebaut werden können: „Selbst bei einer mündlichen Abiturprüfung erfahren sie viel früher, wann sie in die Prüfung gehen.“ (F)

Viele Stimmen bringen aber auch positive Aspekte ein:

- „Ich fand das Losverfahren okay. Die konnten alle ganz gut miteinander. Ich würde da immer argumentieren, man kann es sich auch im späteren Leben nicht aussuchen.“ (F)

- „Ich fand es auch gut, dass die Partner ausgelost wurden. Dadurch haben in der Übungsphase Schülerinnen und Schüler miteinander gesprochen, die vorher noch keine drei Worte miteinander gewechselt haben. Ich habe auch versucht, dass jeder mal mit jedem gesprochen hat“ (E)

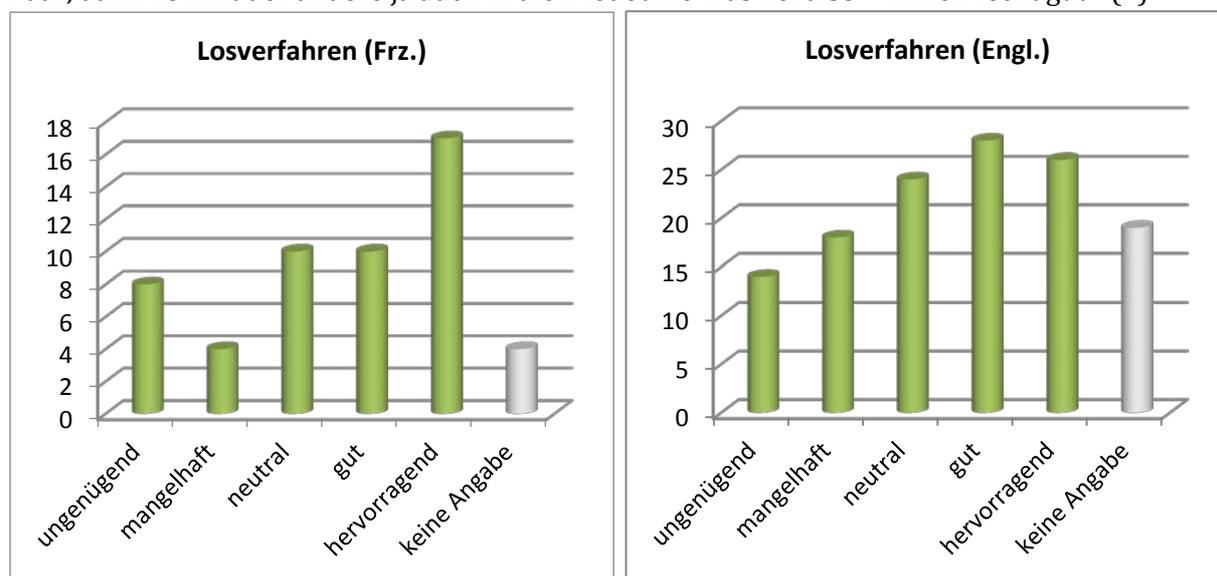
Als weiteres wichtiges Pro-Argument wird die größere Gerechtigkeit des landesweiten Verfahrens angeführt: „Wenn die Schülerinnen und Schüler wissen, alle im Saarland werden nach dem gleichen Prinzip ausgelost, dann gibt es auch Akzeptanz.“ (E)

Andere Lehrende wünschen sich explizit mehr Autonomie in der Zusammenstellung der Prüfungspaare:

- „Ich finde es unklug, wenn Kollegen keinen Einfluss auf dieses Verfahren nehmen können, da es z.B. klug wäre zu sagen, wenn eine Schülerin/ein Schüler hibbelig ist, dass diese/r dann auch als erste/r beginnen kann, da sie/er sonst nur nervöser wird. Oder Paare die einfach nicht funktionieren, da es menschlich, aber auch akustisch nicht geht. Diese Freiheiten sollten den Lehrern gegeben werden.“ (F)
- „Bei einem Fall waren sich die beiden spinnefeind und konnten und wollten nicht. Von daher war ich nicht unbedingt ein Befürworter des Losens.“ (E)
- „Wir hatten einen Totalausfall, aber das war zu erwarten, weil der Schüler auch im Unterricht nichts sagt. Das war für die Partnerin furchtbar. Da hat die Zweitprüferin die Rolle übernommen.“ (E)

Kritisch wird bisweilen die vorgesehene Dreierprüfung zum Ausgleich ungerader Prüflingszahlen gesehen. Hier wird mehrfach darauf hingewiesen, dass diese Konstellation an sich schwerer zu beurteilen ist und ein einzelner Prüfling eventuell nicht adäquat zum Zuge kommt: „Das war schon ein bisschen schwieriger für die, dass da jeder zu Wort gekommen ist. Das hat auch eine Schülerin gesagt, dass sie den anderen nicht das Wort abschneiden wollte und hat dann gewartet, bis sie an der Reihe war und dadurch hat sie am Schluss auch ein bisschen schlechter abgeschnitten.“ (E)

Einzelne Lehrende befürworten die Dreierprüfung hingegen, „weil drei Schüler noch mehr sagen und sich noch mehr den Ball zuspielen können. Wenn es zwei sind und es ist ein schwieriges Paar, dann kommt der andere ja auch in die Bredouille. Das lief also wirklich recht gut.“ (F)



Grafik 35a und 35b: Lehrkräfte - Bewertung des Losverfahrens

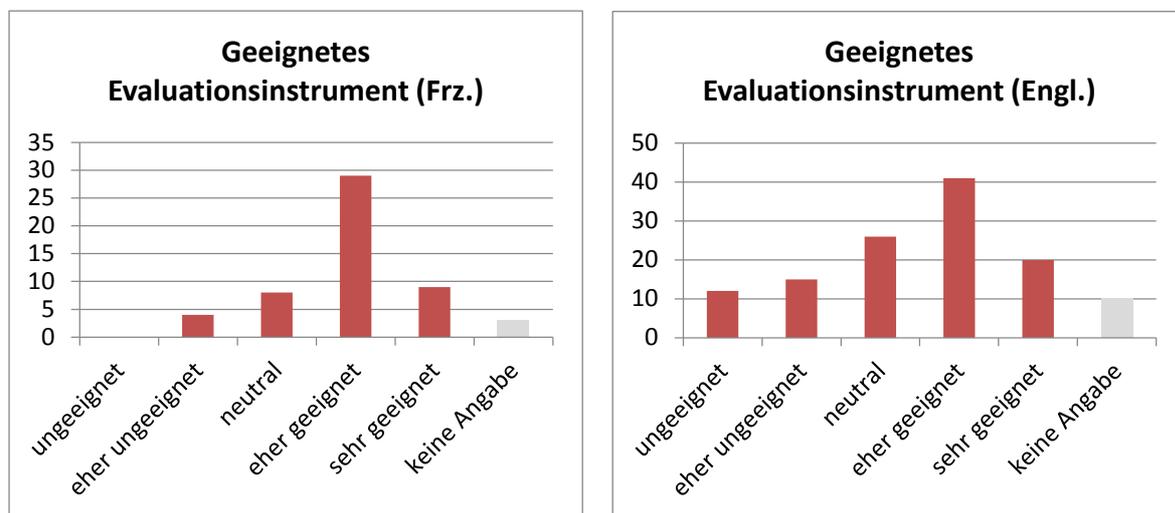
Die landeszentrale Sprechprüfung wird von der überwiegenden Zahl (71,7%) der Französischlehrkräfte als *eher bis sehr geeignetes* Evaluationsinstrument angesehen.

Die Englischlehrkräfte beurteilen dies etwas kontroverser; knapp die Hälfte (49,2%) halten aber auch hier die Sprechprüfung für *eher bis sehr geeignet* (21% sind *neutral* eingestellt und 8,1% machen *keine Angabe*). Viele begrüßen in den Freitextantworten ausdrücklich, dass das

Sprechen mehr Gewicht erhält, und machen ihre Zustimmung von der Optimierung der Organisation abhängig:

- „Sehr gut, wenn nachgebessert wird und Organisationsprobleme gelöst werden“. (E)
- „Es ist die einzige Möglichkeit, objektiv und reliabel die Kompetenz des Sprechens zu evaluieren.“ (E)
- „Positiv: Überprüfung und Bewertung der mündlichen Kommunikation rückt in den Vordergrund, besser als Kursarbeit.“ (E)

Betont wird die „Förderung eines authentischen Sprachgebrauchs, da Sprechen sonst zu kurz kommt.“ (E); „Das ist authentisch, da trifft man auch Menschen, mit denen man vielleicht nicht kann und muss mit ihnen sprechen“ (F).

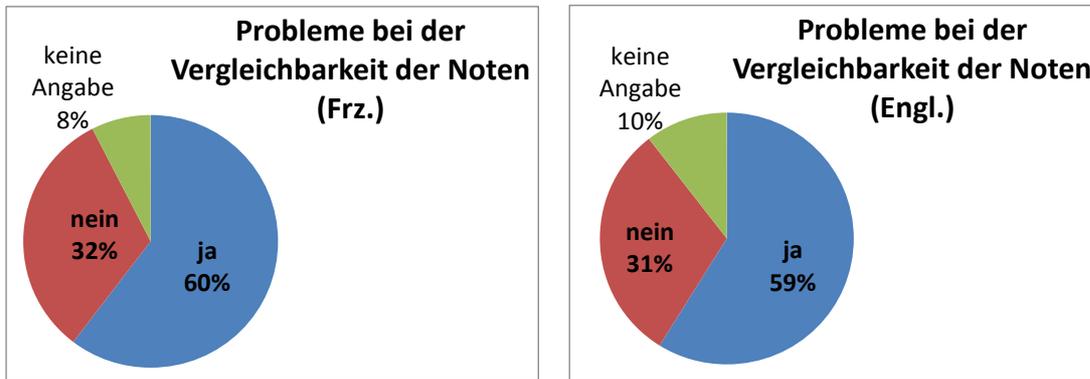


Grafik 36a und 36b: Lehrkräfte - Einschätzung der Sprechprüfungen als geeignetes Evaluationsinstrument

Auch zum Zeitpunkt und der Einbindung der Prüfungen gibt es Vorschläge. So wird angeregt, diese im Zeitraum nach der schriftlichen Abiturprüfung zu terminieren. Bisweilen wird auch die Erwartung geäußert, dass vor dem Hintergrund des organisatorischen Aufwands die Hörverstehensprüfung im Abitur durch die Sprechprüfung ersetzt werden könnte.

Die Frage nach möglichen Problemen bei der Vergleichbarkeit der Noten wird von rund einem Drittel der Französisch- und Englischlehrkräfte verneint, während je rund 60% von ihnen Schwierigkeiten hinsichtlich der Vergleichbarkeit der Noten erkennen. Allerdings wird diese Stellungnahme relativiert durch die ergänzenden Angaben in den Freitexten. Der unterschiedliche Schwierigkeitsgrad und die Gestaltung der Impulse (Anzahl der Impulse, inhaltliche Komplexität), aber auch die fehlende Vorbereitungszeit sowie die Paarzuordnung per Los, werden hier vor allem für eine negative Gesamteinschätzung angeführt: „So schnell [können] nicht [alle] interpretieren.“ (F), oder der Gesprächsstoff sei begrenzt, „sodass die Schülerinnen und Schüler kaum Thema vorgehalten bekommen, sondern selbst etwas finden müssen“. (E)

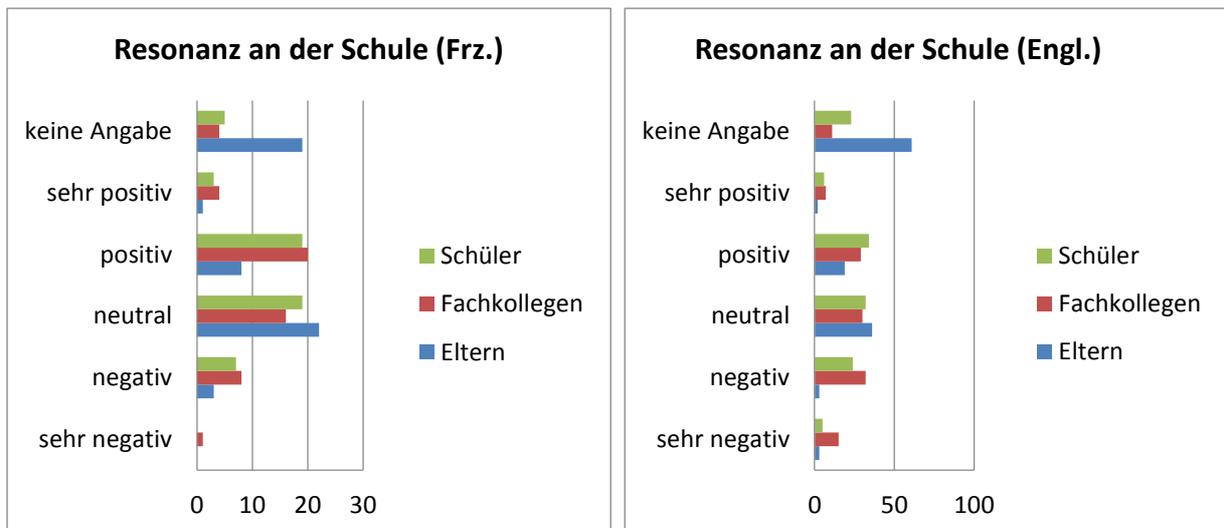
Außerdem beeinträchtigt nach Auffassung vieler Lehrender die Reihenfolge der Prüflinge die Vergleichbarkeit der Noten, wobei Vor- und Nachteile allerdings verschieden verortet werden: „Kandidat A sieht den Impuls, kann direkt sprechen, während B gezwungen ist erst zuzuhören, bei 3er Prüfung kann ein Kandidat übergangen werden“ (E) – „Gibt Partner A schon viel zum Thema vor, wirft man B oftmals vor, die Argumente von A zu wiederholen.“ (E) Es wird auch argumentiert, dass Prüfling B sich mehr Gedanken machen könne; hier wird vorgeschlagen, dass man dies durch mehr Vorbereitungszeit ausgleichen könnte.



Grafik 37a und 37b: Lehrkräfte - Einschätzung der Vergleichbarkeit der Noten

Die Frage nach der Resonanz an den Schulen wird aus Sicht der Lehrenden überwiegend *neutral* bis *positiv* beschrieben; von den Eltern gibt es kaum Rückmeldungen. Die Lehrkräfte sehen das Meinungsbild im Fachkollegium kritischer als bei Eltern und Schülerinnen und Schülern – viele Lehrende haben danach teilweise deutliche Vorbehalte; die stark positiven Meinungen sind weniger ausgeprägt. In den Interviews wird vor allem verstärkt auf die Sichtweise der Schülerinnen und Schüler eingegangen:

- „Die Leute, die die Einsen bekommen, finden das gut. [...] Denjenigen, denen es schwerfällt, die könnten auch auf die Sprechprüfung verzichten.“ (F)
- „Im Prinzip haben sie das akzeptiert. Die haben ein bisschen geschluckt am Anfang, als es um das Losen ging.“ (F) –
- „Und die Franzosen, wir hatten vorher immer einen Teil der monologisch war, den sie ein bisschen vorbereiten konnten, den hatten sie jetzt nicht mehr.“ (F)

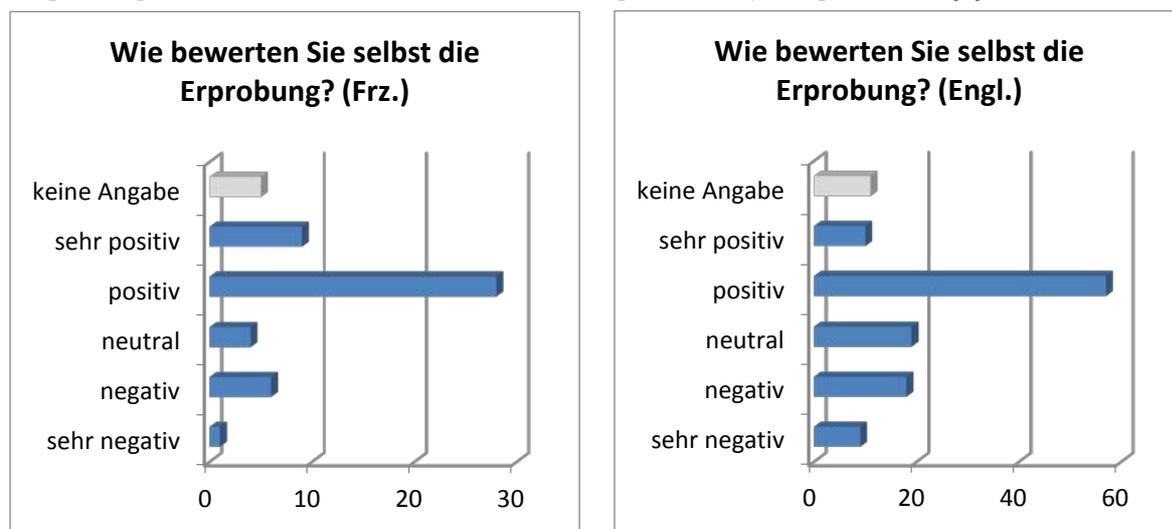


Grafik 38a und 38b: Lehrkräfte - Einschätzung der Resonanz an der Schule

3.2.4 Fazit der Lehrkräfte

Die meisten Lehrenden beurteilen die Erprobung der landeszentralen Sprechprüfung positiv, auch wenn es einige dezidiert negative Stimmen gibt.

„Die Schulleitung hat sich schon sehr viel Mühe gemacht, damit das mit der Organisation klappt. Ansonsten... ich denke auch, dass die Leute, die die Prüfungen zusammengestellt haben, sich viele Mühe gemacht haben, um da wenigstens annähernd gleichwertige Anforderungen zu finden, was trotzdem nicht ganz gelungen ist.“ (F) – „Es gibt ja auch Bundesländer, wo es schon lange Zeit praktiziert wird aber unsere Erfahrungen waren jetzt gemischt.“(F)



Grafik 39a und 39b: Lehrkräfte - persönliche Bewertung der Erprobung

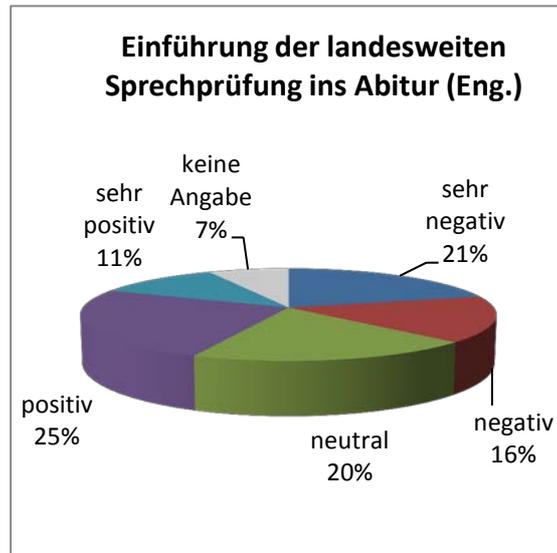
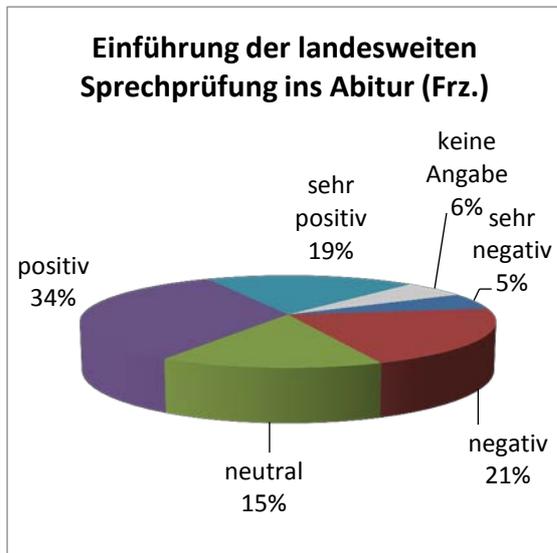
Die geplante Einführung der Sprechprüfungen hingegen polarisiert. Sie wird in einer großen Bandbreite von *sehr negativ* bis *sehr positiv* eingeschätzt, was sich auch in den bereits ausgeführten differenzierten Einzelmeinungen zu verschiedenen Einzelaspekten widerspiegelt. Zwar sieht insgesamt die Mehrheit der Lehrenden die Prüfungen als ein geeignetes Evaluationsinstrument, jedoch ist festzustellen, dass hier die positive Einschätzung von Bedeutung und Sinn der Sprechprüfung in Relation gesetzt wird zum konkret erfahrenen Organisationsaufwand bei der Erprobung: „Ob die Sprechprüfung den Aufwand im Abitur rechtfertigt, das muss man noch mal sehen. Ich denke, dass das von der Zeit und vom Personalaufwand her doch Erhebliches fordert.“ (F) – „An sich eine ganz tolle Sache, aber der Organisationsaufwand viel zu hoch“ (E)

Dabei gibt es insbesondere viele sehr negative Stimmen von Lehrkräften, die eine Einführung entschieden ablehnen. Einige Stimmen von Englischlehrenden:

- „Hoher organisatorischer Aufwand mit kaum neuen Ergebnissen, hohe Belastung im Prüfungsgeschehen, als Abiturprüfung ungeeignet“ (E)
- „Das bisherige Verfahren war absolut ausreichend als Leistungsmessung, die Bewertung durch Schülerinnen und Schüler und Fachkollegium war positiv, die Zentralisierung der Prüfung macht keinen Sinn.“ (E)

Demgegenüber sind viele aufgeführte Kritikpunkte selbst bei negativer Sicht durchaus konstruktiv, und auch bei generell positiven Bewertungen werden sehr konkrete Verbesserungswünsche aufgeführt, wie bei verschiedenen Einzelfragen bereits deutlich wurde (mehr Zeit für Notenfindung und Bewertung sowie für den Einblick in Unterlagen; eine allgemeine Reduzierung der Belastung der Prüfenden, z.B. durch eine Entzerrung des Prüfungstaktes und weitere organisatorische Änderungen⁵).

⁵ Folgende Beispiele aus den Freitextantworten seien exemplarisch genannt: „Durchführung an 2 Vormittagen wünschenswert; Pausen währenddessen könnten länger sein; dialogischer Teil könnte kürzer sein“ (F) – „Frühere Information der Schulen/Prüfer, weniger rigide Prüferskripte, weniger komplexe Protokollbögen, überarbeitete Bewertungsmatrix“ (F) – „Warum keine Vorbereitungszeit von 30 Min, zurückhaltende Schülerinnen und Schüler im Nachteil, warum kein Protokollführer, Losverfahren problematisch“ (F).

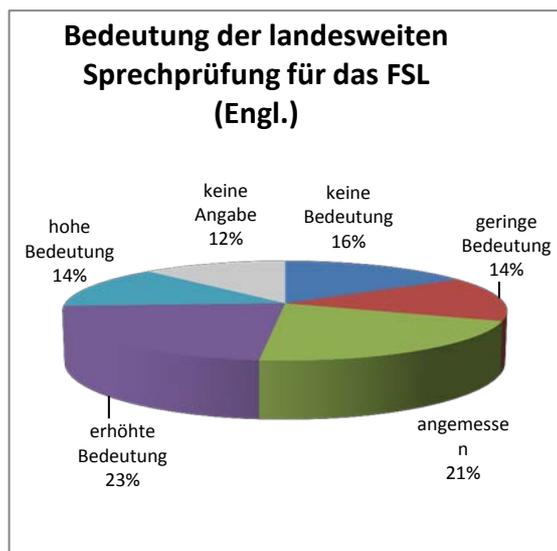
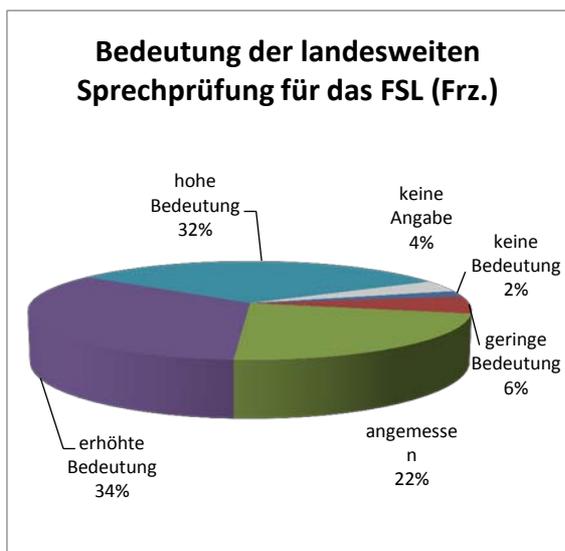


Grafik 40a und 40b: Lehrkräfte - Haltung zur geplanten Einführung der Sprechprüfungen ins Abitur

Trotz dieser differenzierten Haltung wird der landeszentralen Sprechprüfung von den Lehrkräften – ähnlich wie von den Schulleitungen – eine erhebliche Bedeutung für das Fremdsprachenlernen beigemessen, wobei bei den Französischlehrenden die Bewertung insgesamt positiver ausfällt (34% *erhöhte* und 32% *hohe* Bedeutung gegenüber 23% *erhöhte* und 14% *hohe* Bedeutung bei den Englischlehrkräften): „Ich glaube, dass wir die Sprachprüfung an sich sinnvoll finden, da sind wir uns einig.“ (F)

In den Einzelmeinungen und den Interviews wird hervorgehoben, dass die Prüfung der mündlichen Kompetenz eine wichtige und willkommene Bereicherung der Bewertungsmöglichkeiten darstellt, insbesondere wenn die Organisation sich eingespielt hat und die genannten Probleme bearbeitet werden. So führen die Englischlehrkräfte aus:

- „Ich habe mich gefreut, dass endlich eine Prüfung da ist, die die Kompetenzen, die wir anstreben, prüft.“ (E)
- „Ich find’ das von der Idee her gut, dass man spontan redet. Entspricht realen Situationen, die man im Ausland erlebt, da krieg’ ich nicht noch Unterlagen.“ (E)
- „Es war wichtig, dass die sprechen, keine Wissensabfrage. Das Thema war nur ein Anlass, und es waren tolle Themen.“ (E)



Grafik 41a und 41b: Lehrkräfte - Bedeutung der Sprechprüfung für das Fremdsprachenlernen

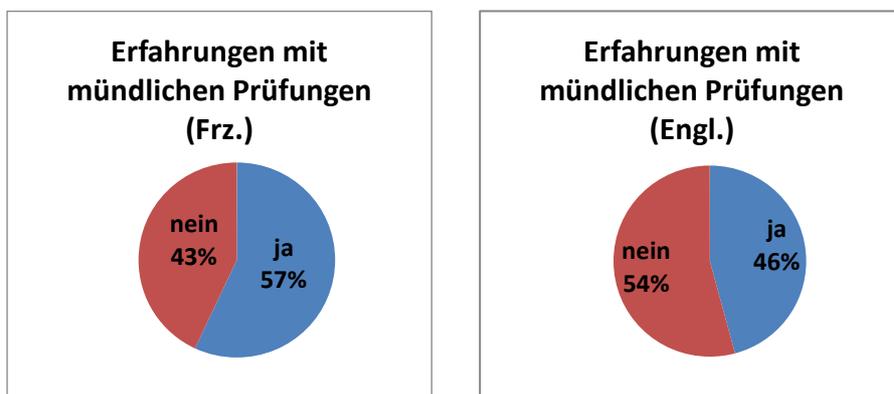
Vor diesem Hintergrund darf insgesamt trotz des differenzierten Meinungsbildes eine grundsätzlich konstruktive Haltung der Lehrenden zur geplanten Einführung der Sprechprüfung konstatiert werden. Teilweise negativen Einschätzungen auf der Basis der Erfahrungen während der Erprobung steht eine erkennbare Wertschätzung für eine gestärkte mündliche Prüfungskomponente gegenüber, wobei die aufgeführten Anregungen zur Optimierung der Prüfungsorganisation im Kontext der konkreten Einführung von Nutzen sein können.

3.3 Schülerinnen und Schüler

3.3.1 Information und Erwartungen

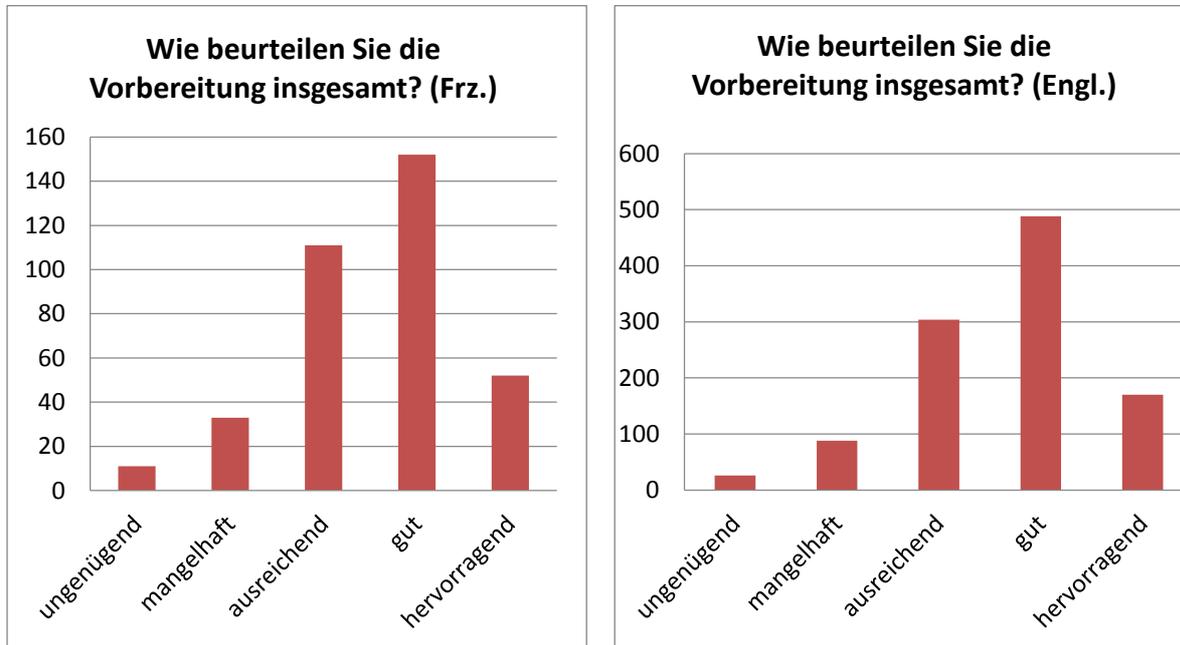
Die befragten Schülerinnen und Schüler wurden durch ihre Fachlehrer über die Durchführung der Sprechprüfung informiert. Sie hatten nach Einschätzung der Lehrkräfte zunächst gleichermaßen positive wie negative Erwartungen an die Prüfung (vgl. Grafik 18a und 18b). Häufig geäußert wurden auf der positiven Seite, mit der Sprechprüfung die eigene Kompetenz besser abbilden zu können; auf der anderen Seite bestand aber auch Unsicherheit insbesondere wegen der Partner- und Themenwahl.

Etwa die Hälfte der Schülerinnen und Schüler gab an, bereits Erfahrungen mit mündlichen Prüfungen gesammelt zu haben; in Französisch waren dies 57%, in Englisch 46%.



Grafik 42a und 42b: Schülerschaft - Erfahrungen mit mündlichen Prüfungen

Die Vorbereitung auf die Prüfungen wurde sowohl für Französisch als auch für Englisch überwiegend als *gut* bewertet (3,56 von 5 für Französisch, 3,64 für Englisch), ebenso die Vorbereitungsmaterialien (3,8 F und E) und die Vorbereitungszeit (3,52 F/ 3,56 E).

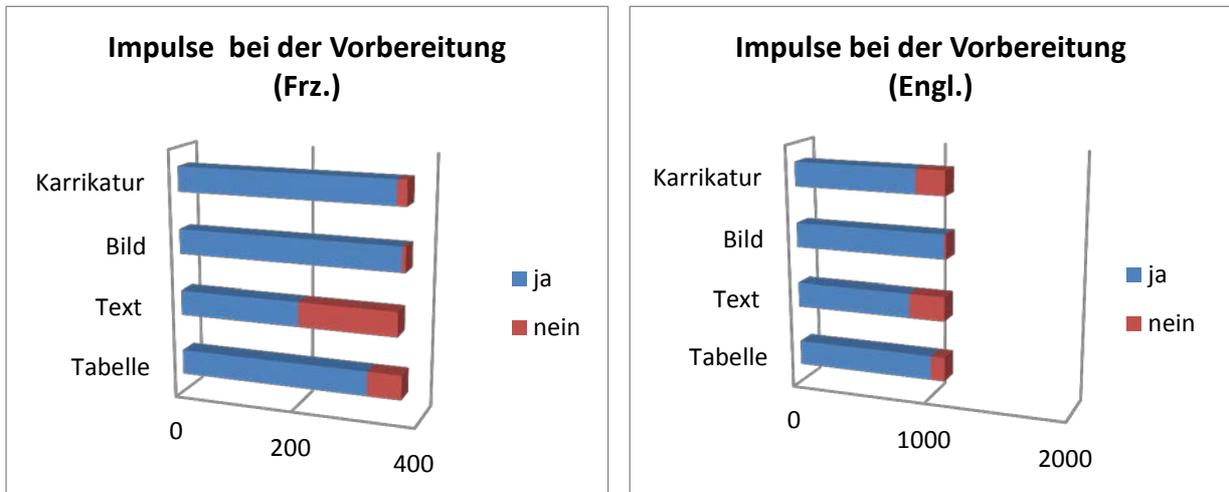


Grafik 43a und 43b: Schülerschaft - Beurteilung der Vorbereitung insgesamt

Die Frage nach der Vorbereitung auf die verschiedenen Impulse im Vorfeld der Prüfungen wurde von den Schülerinnen und Schülern überwiegend bejaht, wobei der Fokus allerdings

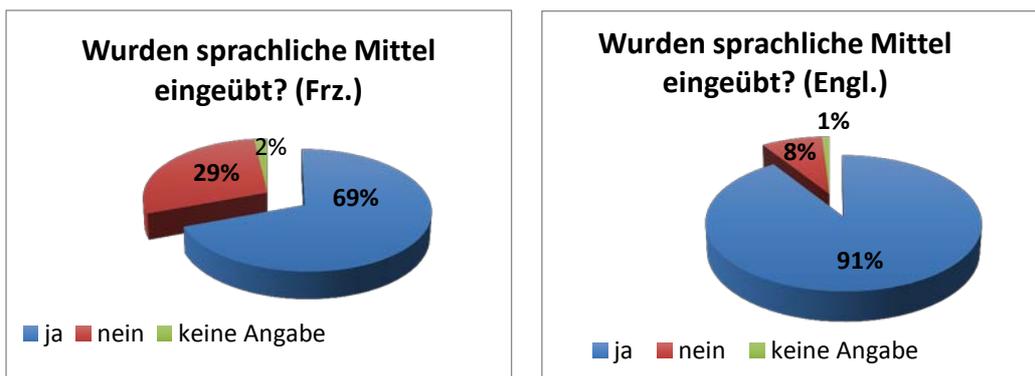
deutlich auf Bild und Karrikatur lag. Nach den Angaben der Schülerinnen und Schüler wurden nicht überall alle Prüfungsformen eingeübt.

Während nahezu alle Schülerinnen und Schüler Bildimpulse geübt haben, lag der Vorbereitungsgrad nach den Angaben der Schülerinnen und Schüler in Französisch bei Karikaturen (95,7%), Tabellen (84,7%) und Texten (54%) und, in Englisch bei Tabellen (89,7%), Karikaturen (79,9%) und Texten (75,5%).



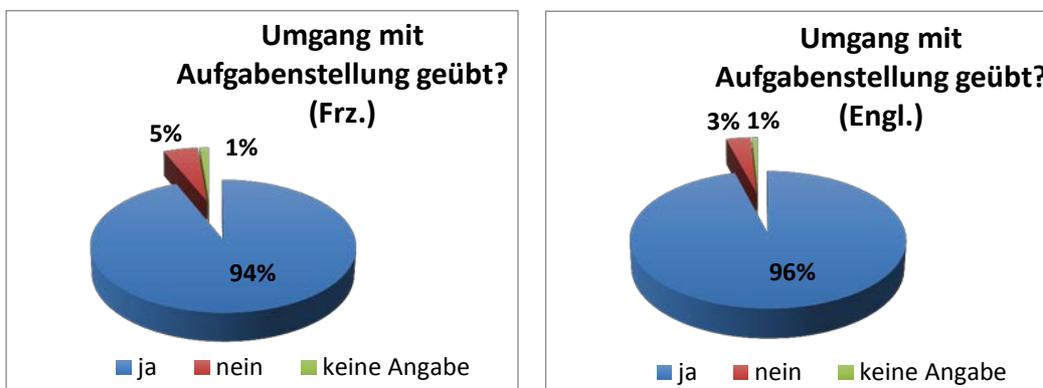
Grafik 44a und 44b: Schülerschaft - Impulse bei der Vorbereitung

Während in Englisch 91% der Schülerinnen und Schüler angaben, sprachliche Mittel im Unterricht geübt zu haben, waren es in Französisch nur 69%.



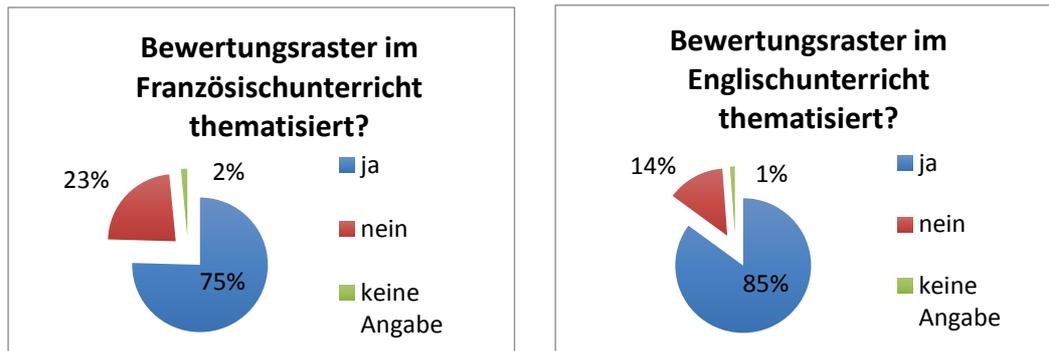
Grafik 45a und 45b: Schülerschaft - Einüben sprachlicher Mittel im Unterricht

Bis auf 5% (Französisch) bzw. 3% (Englisch) haben alle Schülerinnen und Schüler die Aufgabenstellungen anhand der zur Verfügung gestellten Probetests geübt.



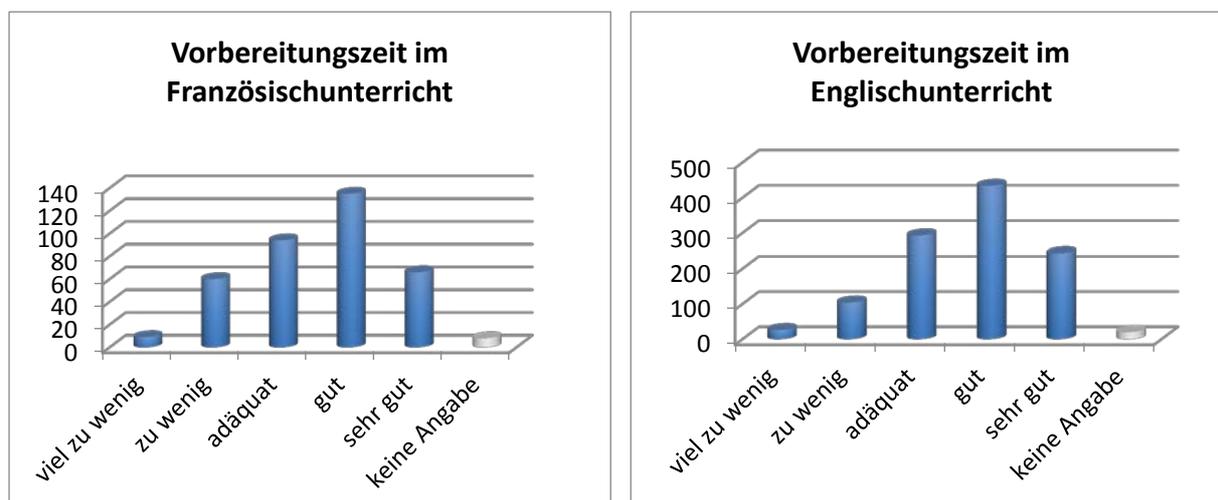
Grafik 46a und 46b: Schülerschaft - Umgang mit Aufgabenstellungen

Die Bewertungsraster wurden nach Angaben von drei Vierteln der Französischschülerinnen und -schüler und 85% der Englischschülerinnen und -schüler im Unterricht thematisiert.



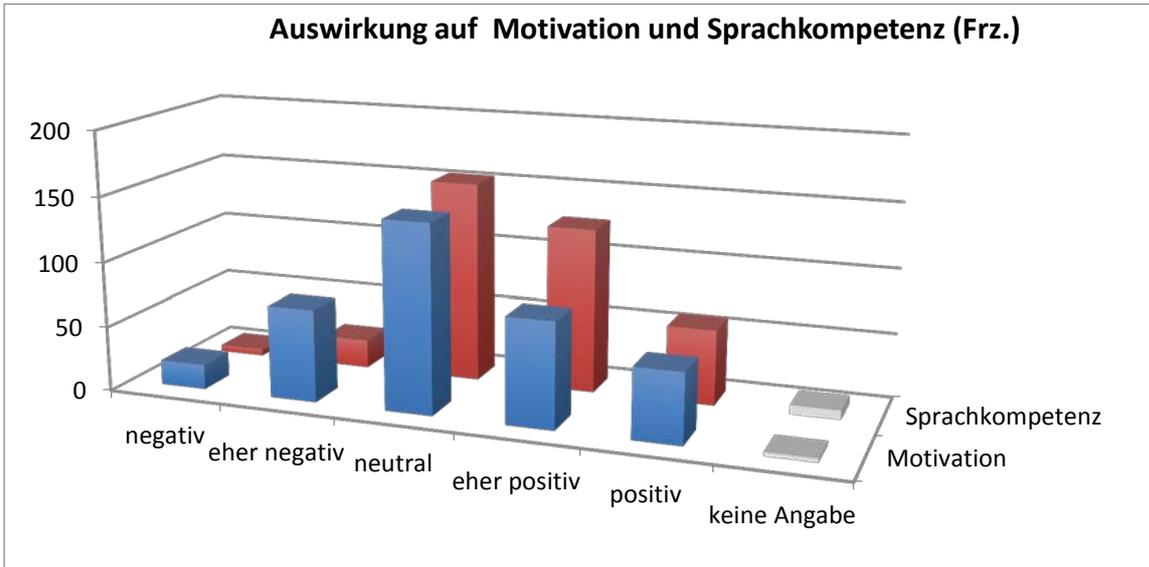
Grafik 47a und 47b: Schülerschaft - Umgang mit Bewertungsraster

Die im Unterricht zur Verfügung gestellte Vorbereitungszeit wird von den befragten Schülerinnen und Schülern als überwiegend *adäquat* bis *sehr gut* bewertet. Die Verteilung liegt im Einzelnen für Englisch bei 26% und für Französisch bei 25% *adäquat*, 39% (E) bzw. 36% (F) *gut*, 22% (E) bzw. 18% (F) *sehr gut*. Insgesamt liegen also 87% (E) bzw. 79% (F) positive Rückmeldungen vor, wobei allerdings auch von einem nicht geringen Anteil (11,6% in Englisch und 18,5% in Französisch) zu wenig Vorbereitungszeit geltend gemacht wird.

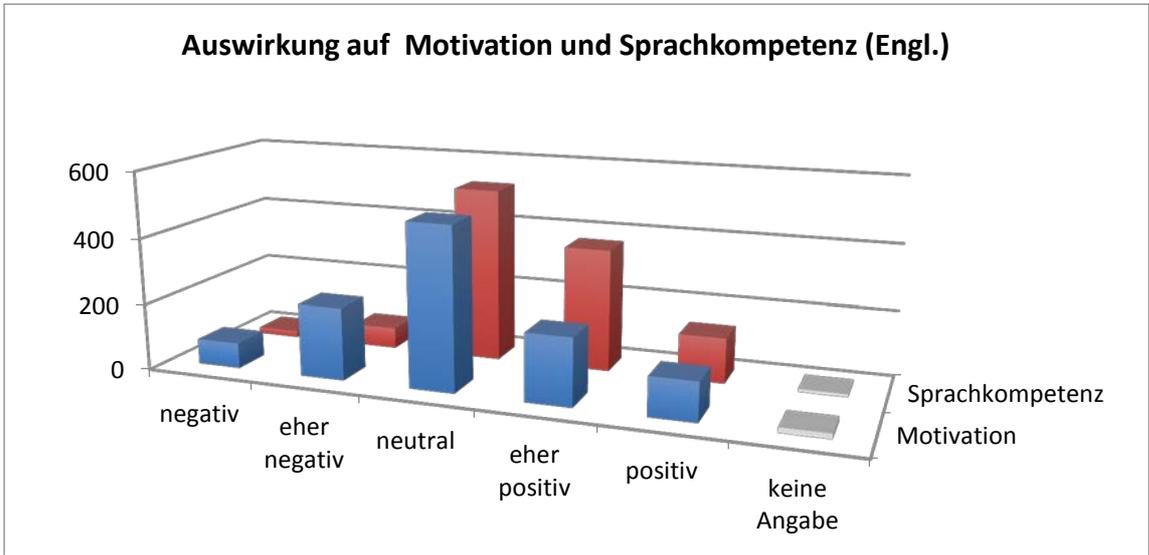


Grafik 48a und 48b: Schülerschaft - Vorbereitungszeit im Unterricht

Nach Ansicht der Schülerinnen und Schüler hat sich die Sprechprüfung insgesamt eher positiv auf ihre mündliche Kompetenz ausgewirkt (im Mittel 3,58 von 5 in Französisch; 3,49 in Englisch). In Bezug auf die eigene Motivation fällt die Antwort für Englisch etwas weniger positiv aus (im Mittel 3,22 von 5 in Französisch; 3,06 in Englisch); hier halten sich negative (27%) und positive (28%) Einschätzungen die Waage; die meisten Schülerinnen und Schüler sehen die Auswirkungen der Sprechprüfung auf die eigene Motivation eher neutral (44%).

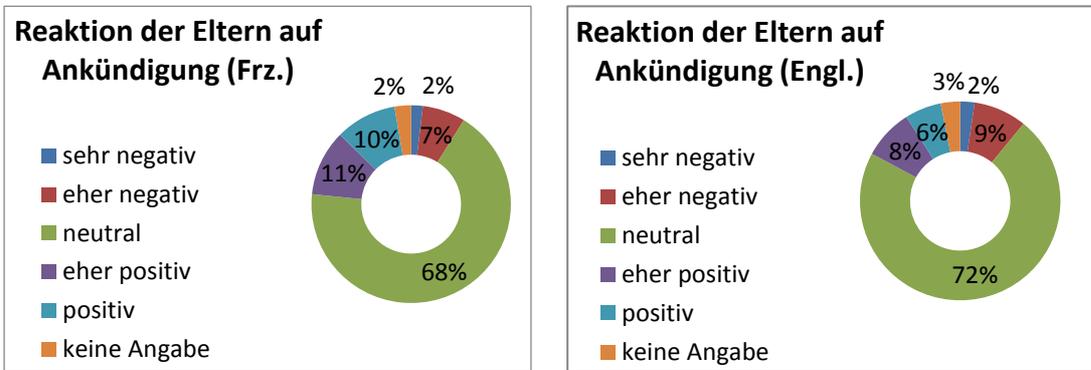


Grafik 49a: Schülerschaft - Auswirkung auf Motivation und Sprachkompetenz Französisch



Grafik 49b: Schülerschaft - Auswirkung auf Motivation und Sprachkompetenz Englisch

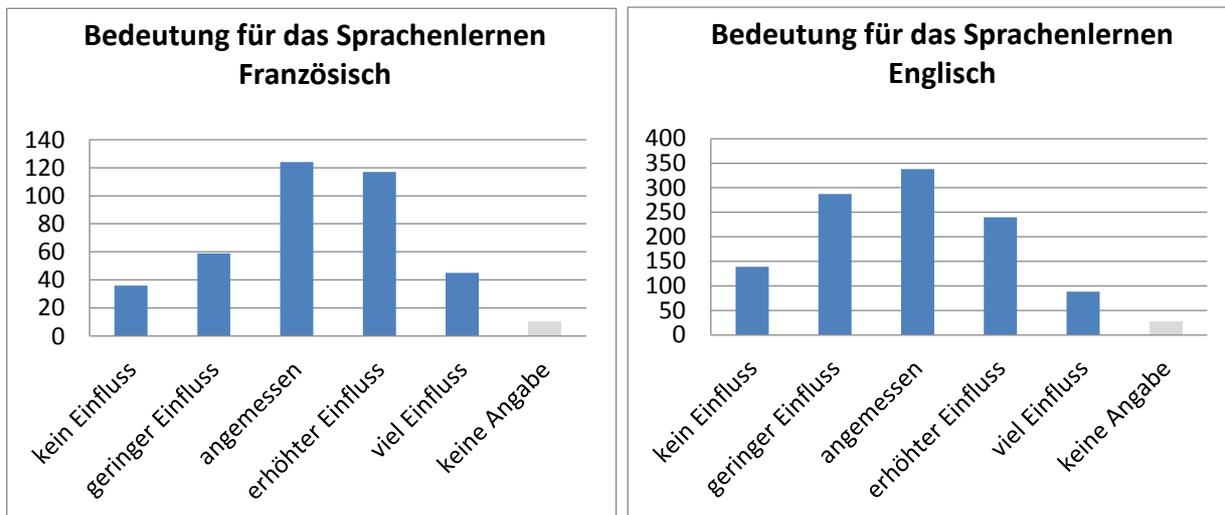
Die Reaktion der Eltern auf die Ankündigung der Sprechprüfung wird von den Schülerinnen und Schülern überwiegend (68% Französisch, 72% Englisch) mit neutral angegeben. Wenigen eher negativen bzw. negativen Reaktionen (17% Französisch, 11% Englisch) stehen 21% (Französisch) bzw. 14% (Englisch) eher positive und positive Einschätzungen gegenüber.



Grafik 50a und 50b: Schülerschaft - Reaktion der Eltern auf die Ankündigung der Prüfung

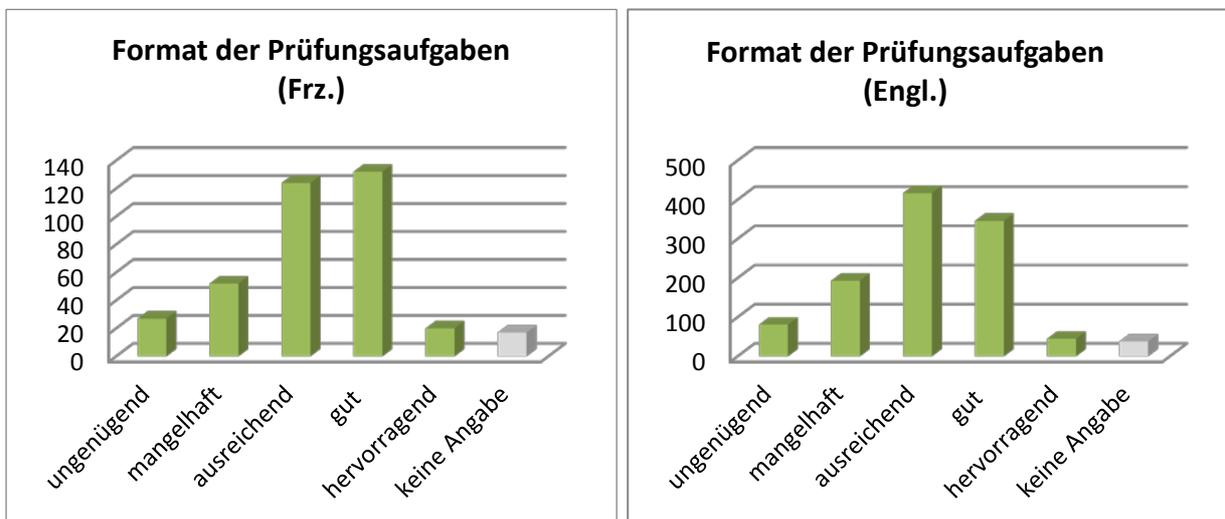
3.3.2 Prüfungsformat und Schwierigkeit

Die Schülerinnen und Schüler messen der Sprechprüfung mehrheitlich (59% für Englisch und 77% für Französisch) einen *angemessenen*, *erhöhten* oder *großen* Einfluss auf das Fremdsprachenlernen bei, etwa ein Viertel (Französisch 26%) bzw. etwas mehr als ein Drittel (Englisch 38%) allerdings nur einen *geringen* bis *keinen* Einfluss:



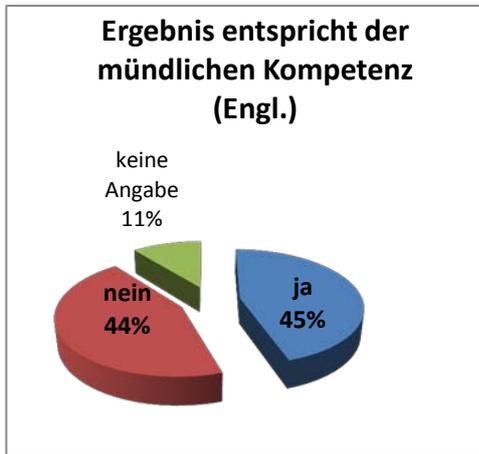
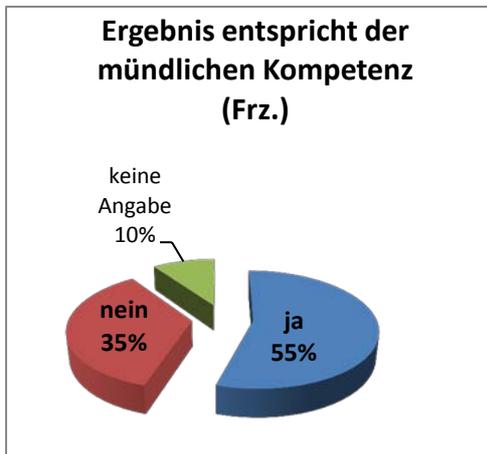
Grafik 51a und 51b: Schülerschaft - Bedeutung für das Sprachenlernen

Die Formate der Prüfungsaufgaben finden ein kontroverses Echo. Von einem Drittel der Französischschülerinnen und -schüler (33%) und 37% der Englischschülerinnen und -schüler werden sie als *ausreichend* bewertet, etwas mehr als ein weiteres Drittel (Englisch 35%, deutlicher Französisch mit 41%) schätzen sie als *gut* oder *hervorragend* ein. Etwa ein Fünftel der Französischschülerinnen und -schüler (21%) und ein Viertel (24%) der Englischschülerinnen und -schüler bewerten sie hingegen als *mangelhaft* oder *ungenügend*.



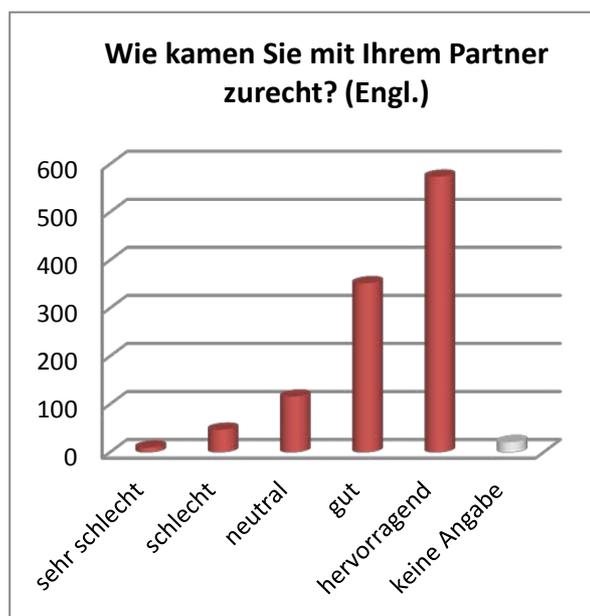
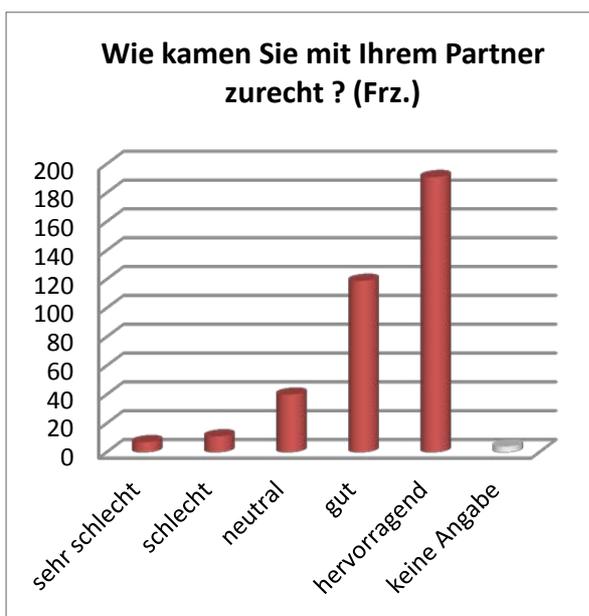
Grafik 52a und 52b: Schülerschaft - Format der Prüfungsaufgaben

Das Ergebnis der landeszentralen Sprechprüfung spiegelt nach Ansicht von 54% der Französisch- und 45% der Englischschülerinnen und -schüler ihre mündlichen Kompetenzen wider, wobei der hohe Anteil derer, die hier keine Angaben machen (10% in Französisch, 11% in Englisch), darauf zurückzuführen sein könnte, dass zum Zeitpunkt der Befragung noch nicht alle Prüflinge ihr Ergebnis kannten:



Grafik 53a und 53b: Schülerschaft - Ergebnis der Sprechprüfung und mündliche Kompetenz

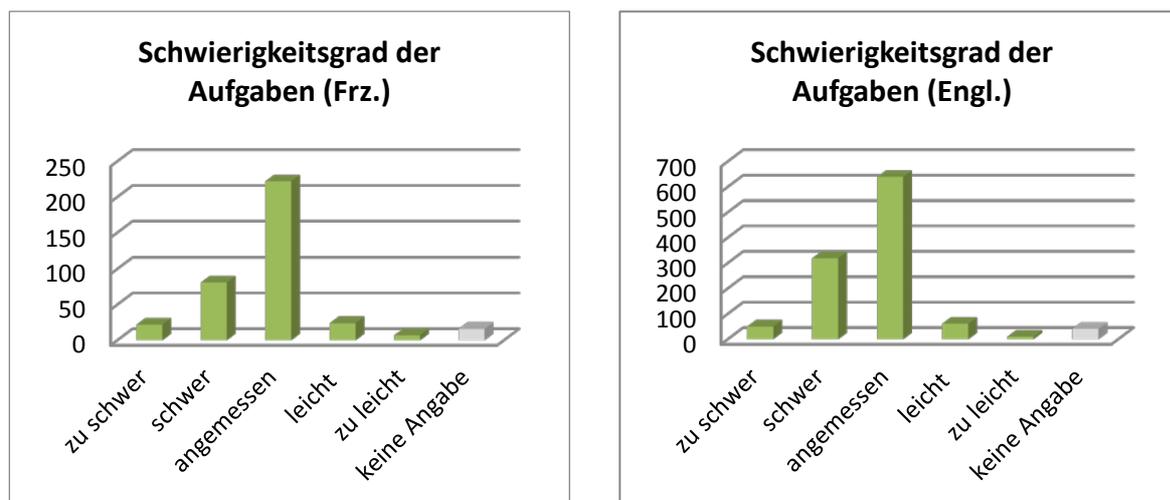
Die Frage, wie sie mit ihrem Partner zurechtkamen, beantworten die Schülerinnen und Schüler ganz deutlich positiv, mehr als die Hälfte (Französisch und Englisch 51%) sogar mit *hervorragend*; der Mittelwert liegt für Englisch wie Französisch bei 4,3 von 5. *Schlecht* oder *sehr schlecht* bewerten nur 5% der Englisch- und 4,8% der Französischschülerinnen und -schüler ihre Interaktion mit dem Partner. Die beiden letztgenannten Ergebnisse führen zu dem Schluss, dass die zufällige Zuteilung der Paare durch das Losverfahren in diesem Kontext grundsätzlich keine negativen Auswirkungen hatte.



Grafik 54a und 54b: Schülerschaft - Wie kamen Sie mit Ihrem Partner zurecht?

Diese Ergebnisse werden in Kap.3.4 zusätzlich im Lichte der Auswertungen zu den beiden Schulen, die nicht am Losverfahren teilgenommen haben, interpretiert.

Der Schwierigkeitsgrad der Aufgaben wird meist als *angemessen* bezeichnet (2,76 für beide Sprachen); die vergleichsweise wenigen negativ urteilenden Schülerinnen und Schüler geben oft an, im Unterricht andere Themen behandelt zu haben bzw. weisen auf den unterschiedlichen Schwierigkeitsgrad der verschiedenen Prüfungsaufgaben hin:



Grafik 55a und 55b: Schülerschaft - Schwierigkeitsgrad der Aufgaben

Die Vergleichbarkeit der Prüfungsimpulse wird von einer Mehrheit (59% Französisch und 60% Englisch) der Schülerinnen und Schüler bejaht. Dieser Befund entspricht im Wesentlichen der Lehrersicht (vgl. Grafik 22a und 22b), die allerdings etwas weniger deutlich ausfällt. Diejenigen Schülerinnen und Schüler, die dies verneinen (38% Englisch und Französisch) nennen in den Freitextantworten verschiedene Gründe dafür. Sie weisen vor allem auf grundlegende Unterschiede zwischen den verschiedenen Impulstypen hin: „Große Unterschiede bei den Impulsen und dem Schwierigkeitsgrad (z.B. sehr schwere Karikatur).“ (E)

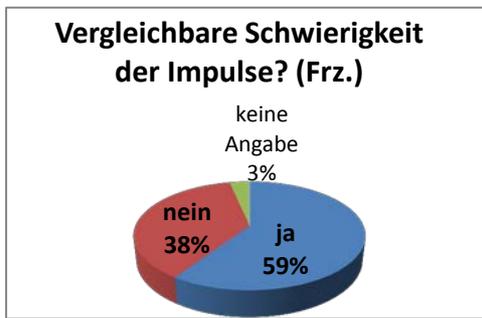
Dabei werden Bildimpulse größtenteils als am leichtesten eingestuft. Texte und Tabellen benötigen dagegen nach Auffassung der Schülerinnen und Schüler mehr Vorbereitungszeit und stehen meist am anderen Ende der Schwierigkeitsskala. Karikaturen werden häufig als problematisch, wenn auch leichter als Tabellen bewertet. Diese werden manchmal nicht verstanden: Die „Intention einer Karikatur kann eher falsch verstanden werden als bei Texten“ (E), so ein Schüler.

Anderer weisen aber auch darauf hin, dass Karikaturen manchmal leichter zu beschreiben seien: „Es ist leichter über eine Karikatur oder ein Bild zu sprechen als eine Stellenbeschreibung.“ (E) Oder: „Erhält man ein Bild ist es leichter einen Einstieg zu finden. Bei Text oder Tabelle ist dies deutlich schwerer.“ (E) Andererseits gebe der „Text [...] Vokabeln vor, die man beim Monolog aufgreifen kann, Bildimpuls bietet diese Hilfestellung nicht.“ (E)

Viele Schülerinnen und Schüler bemängeln, dass sie (mit einer Minute) zu wenig Vorbereitungszeit zur Erfassung der Impulse hatten. Generell werden manche Themen als „wenig ergiebig“ (E) und „nicht interessant“ (E) bemängelt - manche Bilder oder Karikaturen böten nicht genug Redeanlass, „um 3 Minuten zu füllen“ (E), während „Fachvokabular zu Politik und Umwelt unbekannt“ sei, „da Themen nicht interessieren“ (E), wie ein Schüler anmerkt.

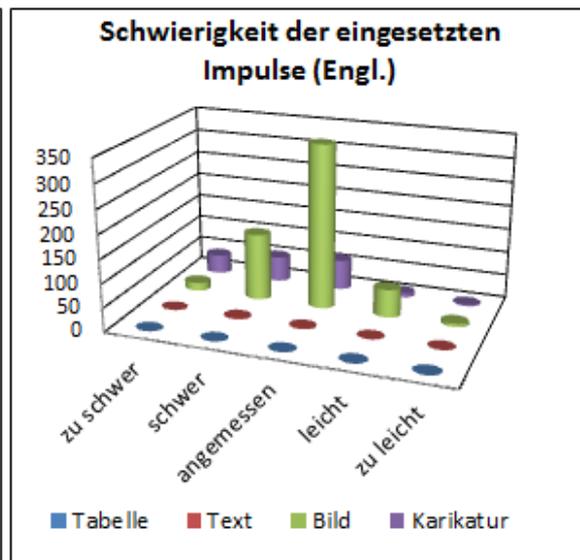
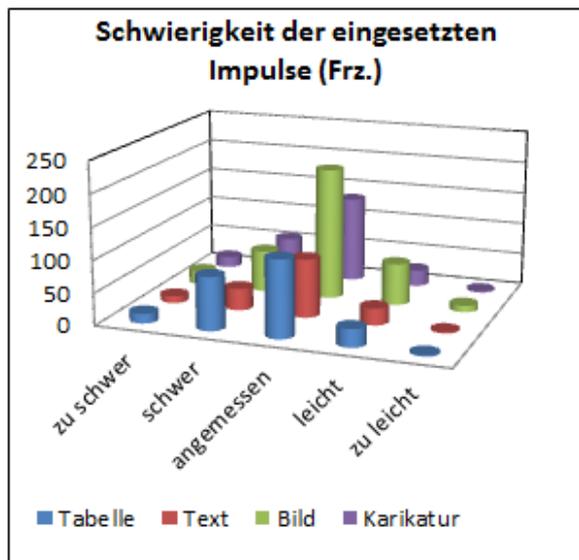
Ebenso wird darauf hingewiesen, dass auch innerhalb der Impulskategorien die Schwierigkeit variiert, was sich z.B. in Unterschieden im Schwierigkeitsgrad zwischen Prüfling A und B innerhalb einer Prüfung niederschlagen kann: „Partner A: kaum beschreibbare Bilder, kaum link zu Aufgabenstellung; Partner B: direkter Impuls zu modern media.“ (E)

Einige Schülerinnen und Schüler merken auch an, dass die Vorbereitung für die Prüfung schwerere Impulse beinhaltet, während die Prüfung selbst leichtere Themen bot (vgl. auch Kap. 3.2.3.2 der Lehrerauswertung).



Grafik 56a und 56b: Schülerschaft - Vergleichbare Schwierigkeit der Impulse

Auch in der Bewertung der jeweils zugeteilten Aufgabe spiegelt sich das Empfinden unterschiedlicher Schwierigkeitsgrade wider; so stufen die Französischschüler, die Bildimpulse erhielten, diese ganz überwiegend als *angemessen* ein, während der Tabellenimpuls häufig als *schwer*, vereinzelt auch *zu schwer* bezeichnet wird:



Grafik 57a und 57b: Schülerschaft - Schwierigkeit der eingesetzten Impulse

Die Gestaltung der Prüfungsmaterialien wird insgesamt überwiegend positiv beurteilt. Dies gilt sowohl für *Prise de Contact/ Warm-up* (3,48 Französisch/3,46 Englisch von 5) als auch für monologischen (3,39 Französisch/3,27 Englisch von 5) und dialogischen Teil (3,63 Französisch/3,60 Englisch von 5). In den Freitextantworten wird sehr oft gelobt, dass die Materialien und die Thematik gut zu erfassen seien. Die Bilder seien klar und gut erkennbar, die Gestaltung übersichtlich, die Aufgabenstellung verständlich und klar formuliert. Hier einige Einzelstimmen von Schülerinnen und Schülern:

- „Prüfungsmaterial übersichtlich aufgebaut, Schwerpunkt und wichtige Schlagwörter auch in knapper Zeit erkennbar“ (E)
- „Prüfungsmaterial gut und übersichtlich gestaltet“ (F)
- „Meist sehr ansprechend, Bilder in Farbe, so dass alles problemlos zu erkennen war“ (E)
- „Große bunte Bilder, alles sehr übersichtlich“ (E)
- „Sie waren farbig, toll! Schriftart war schön. Kompliment!“ – „Kreative, greifbare Aufgabenstellung“ (E)

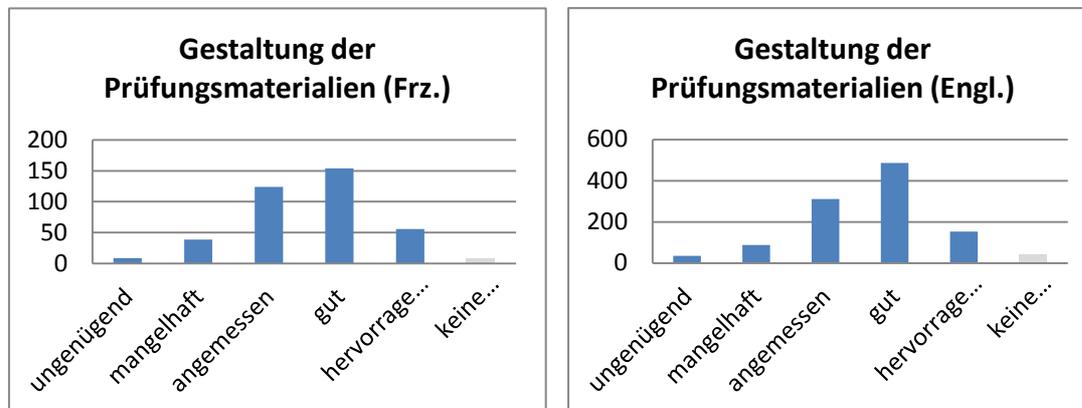
Negativ wird hingegen wiederholt angeführt, dass zu wenig Zeit vorhanden gewesen sei, um die Impulse vollständig zu erfassen; so merkt ein Schüler an: „Aufgrund geringer Vorbereitungszeit unübersichtlich (Anfangs Konzentration auf 1. Aufgabenteil, dabei vergisst man Aufgabenstellung des 2. Teils)“ (E).

Weiterhin hätten einige Impulse zu wenig Aussagekraft, um viel darüber zu sagen:

- „Man konnte nicht viel über die Bilder sprechen“ (E)

- „Bildern fehlt es an Details, womit man 8 min. füllen kann.“ (E)
- „Bilder, auf denen wenig zu beschreiben war“ (F)
- „Bilder nicht aussagekräftig genug“ (F)

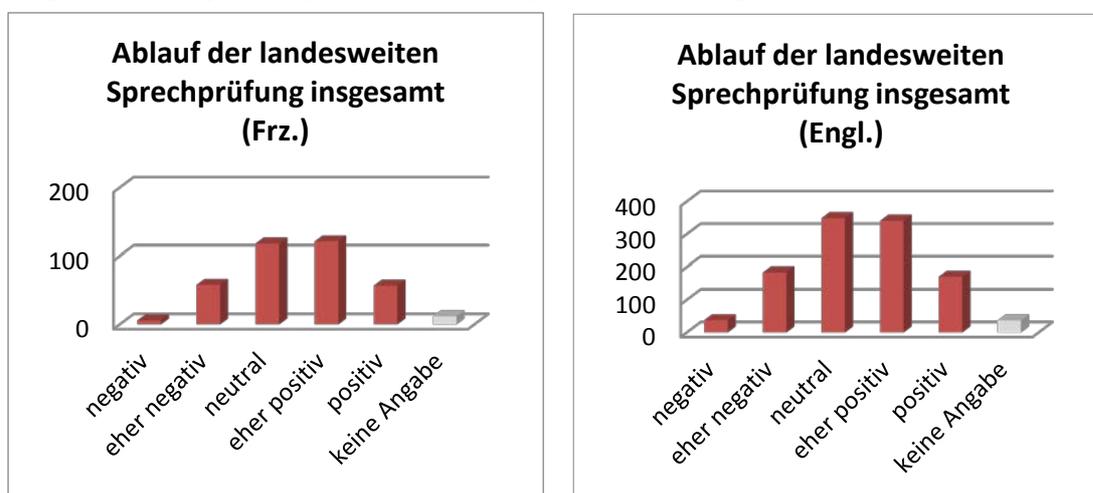
Der Impuls Karikatur wird kontrovers eingestuft: „Die Bilder waren gut zu erkennen, Karikatur jedoch schwer in der Zeit zu deuten.“ (E) Einige Schülerinnen und Schüler bemängeln unklare Aufgabenstellungen, meist im Kontext mit als unklar empfundenen Bildern: „Übersichtlich; wusste genau, was zu machen war; jedoch einige Details in Bild schwer erkennbar“ – „In Bilder konnte man nicht so viel reininterpretieren.“ (E)



Grafik 58a und 58b: Schülerschaft - Gestaltung der Prüfungsmaterialien

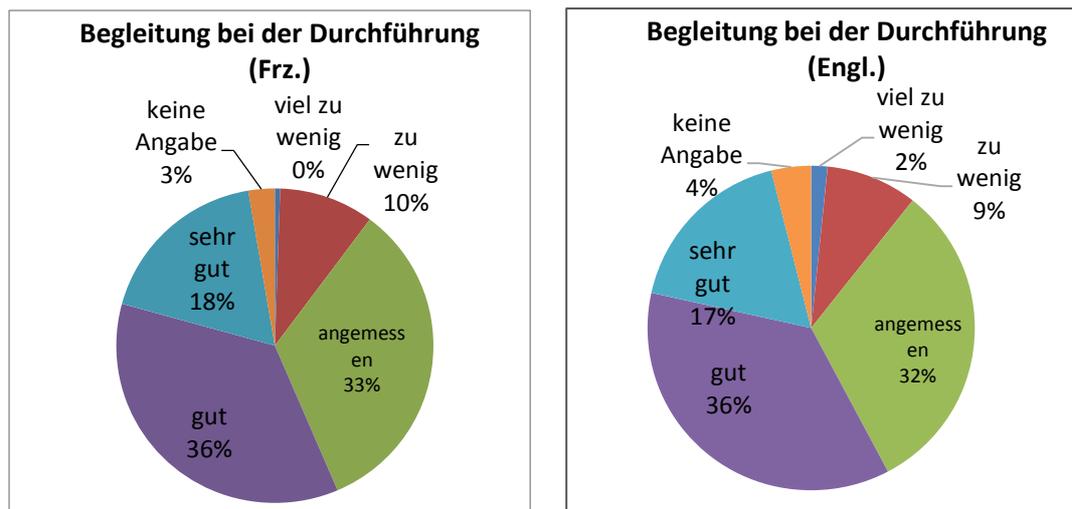
3.3.3 Prüfungsablauf und Benotung

Der Prüfungsablauf wird von den Schülerinnen und Schülern mit 3,46 von 5 (Französisch) und 3,43 von 5 (Englisch) insgesamt überwiegend positiv bewertet. Von den negativ bewertenden Schülerinnen und Schüler werden als Gründe häufig Zeitprobleme (zu lange Wartezeit oder zu geringe Vorbereitungszeit) genannt, weniger das Thema. Zu lange Wartezeit oder wenig Vorbereitungszeit werden auch von Schülerinnen und Schülern, die die Prüfung insgesamt eher positiv (zwischen 4 und 5) bewerteten, angeführt.



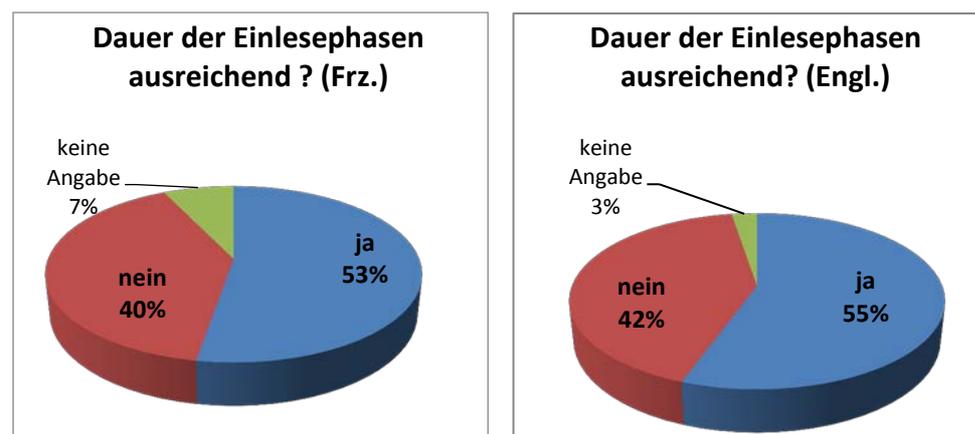
Grafik 59a und 59b: Schülerschaft - Ablauf der landesweiten Sprechprüfung insgesamt

Mit 3,63 von 5 fühlten sich die Französischschülerinnen und -schüler gut begleitet bei der Durchführung, ebenso die Englischschülerinnen und -schüler (3,62).



Grafik 60a und 60b: Schülerschaft – Begleitung bei der Durchführung

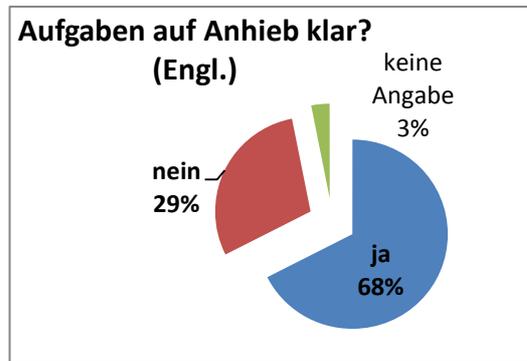
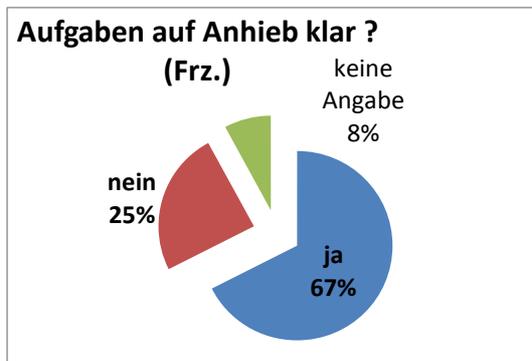
Die Dauer der Einlesephase wird zwar von mehr als der Hälfte der Französisch- (53%) und der Englischschülerinnen und -schüler (55%) als *ausreichend* bezeichnet, doch je etwa 40% und damit fast die andere Hälfte gibt demgegenüber hier an, dass sie nicht ausreichend war. Auch hier zeigen sich Übereinstimmungen mit den Einschätzungen der Lehrkräfte (vgl. Kap. 3.2.3.2)



Grafik 61a und 61b: Schülerschaft - Dauer der Einlesephasen ausreichend?

Für mehr als zwei Drittel der Schülerinnen und Schüler war die Aufgabenstellung auf Anhieb klar (71% der Französisch- und 68% der Englischschülerinnen und -schüler), was sich weitgehend mit der Sicht der Lehrkräfte deckt (Grafik 25a und 25b). Im Freitextfeld differenzierten die gut 30%, bei denen dies nicht der Fall war, die möglichen Gründe dafür. Hier wurde erneut hauptsächlich moniert, dass im Rahmen des Einlesens nicht genug Zeit gelassen wurde, um die Aufgabenstellung zu verstehen und sich vorzubereiten. Einige Einzelstimmen mögen dies illustrieren:

- „Man kann nicht lang genug über das Thema nachdenken. Seriously - 1 min?“ (E)
- „Zu wenig Zeit Aufgabenstellung zu lesen, vor dem Dialog wäre kleine Pause notwendig“ (E)
- „Zu wenig Zeit um sich genügend Gedanken über die Aufgabenstellung zu machen, dann musste man ab und zu stottern.“ (E)



Grafik 62a und 62b: Schülerschaft - Aufgaben auf Anhieb klar?

Außerdem werden auch eine unklare Aufgabenstellung an sich, ein mangelnder Bezug zum Thema, unklare oder uninteressante Themen, nichtssagende Bilder und verwirrende Dialoganweisungen (meist aber bei klaren Operatoren) genannt:

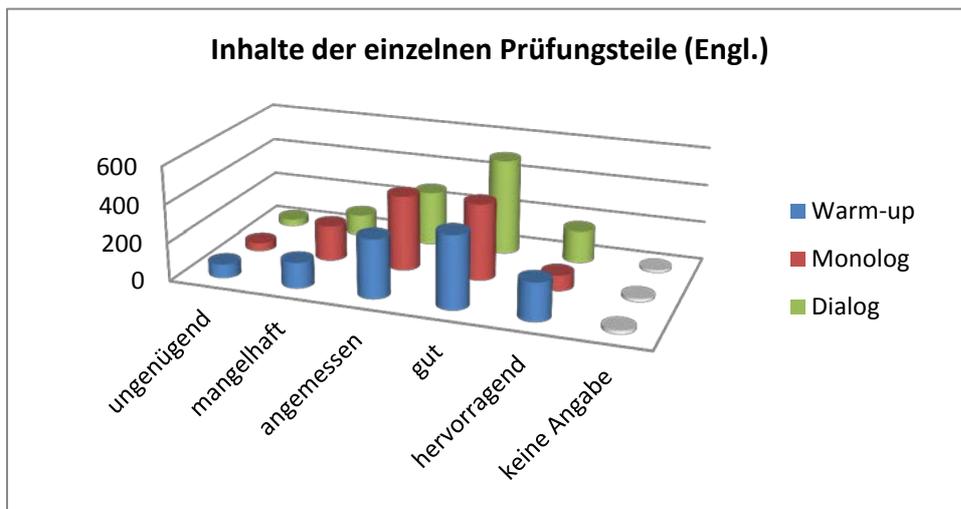
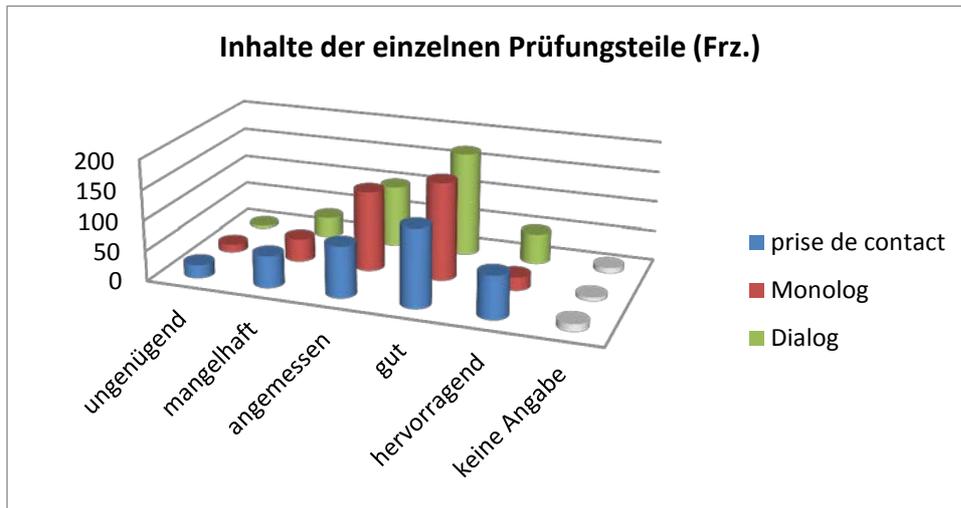
- „Nicht klar deutlich, was genau gemeint, leicht falsch zu verstehen, Aufgabenstellung auf ersten Blick nicht klar erkennbar“ (E)
- „Bei manchen Karikaturen wurde nicht direkt klar, welches Thema bearbeitet werden muss.“ (E)
- „Schweres Thema, was nicht auf Anhieb deutlich war.“ (E)
- „Schweres Verständnis des Bildes/der Karikatur, Aufgabenstellung war unklar“ (F)

Einige Schülerinnen und Schüler thematisieren auch unpassende Operatoren: „Operator compare and contrast unpassend zum Bild, man sollte hier vielleicht weniger steif an den Operatoren festhalten und eine passende Frage stellen.“ (E)

An schülerabhängigen Faktoren werden eigene Probleme beim Verstehen der Aufgabe, Aufregung und unbekannte Vokabeln aufgeführt.

- „Unbekannte Vokabel, ansonsten gute Fragestellung, gutes Material“ (E)
- „(Vor Aufregung) einen Teil der Aufgabe vergessen“ (E)
- „Unbekannte Vokabeln, Operatoren waren aber verständlich.“ (E)
- „Aufgabenstellung durch fehlende Vokabeln etwas schwer verständlich“ (F)

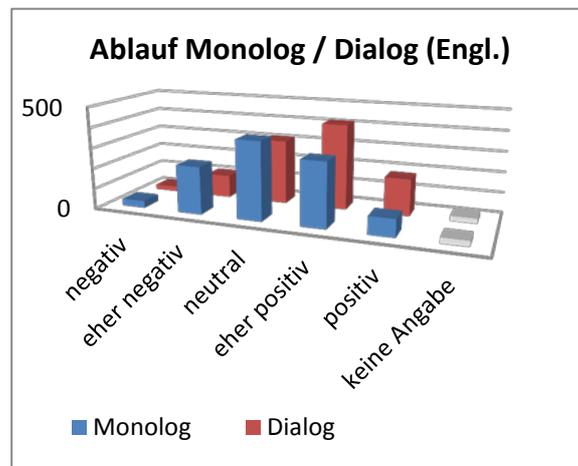
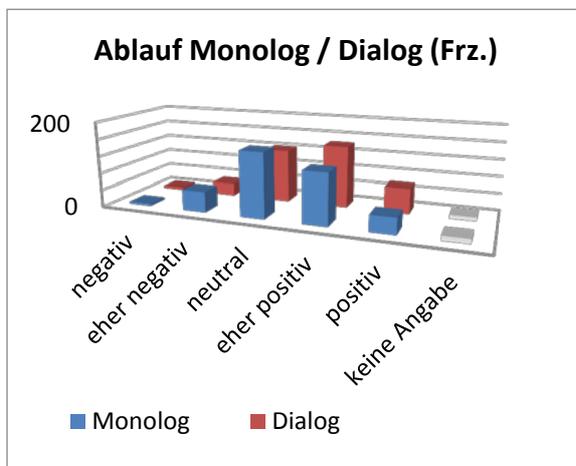
Hinsichtlich der Prüfungsinhalte werden die verschiedenen Prüfungsteile insgesamt überwiegend positiv bewertet, ebenso die Abläufe der verschiedenen Prüfungsabschnitte, wobei der dialogische Teil die höchsten Werte erhält, der monologische die niedrigsten. Im Fach Englisch liegen die Mittelwerte für die Inhalte im Warm-up-Teil bei 3,46 von 5, im monologischen Teil bei 3,27 und im dialogischen Teil bei 3,6. Die Französischschülerinnen und Schülern geben hierfür exakt denselben Wert an; die Mittelwerte für die Inhalte der *Prise de contact* liegen bei ihnen bei 3,48 von 5 und für den monologischen Teil bei 3,39. Die Einschätzung der Prüfungsinhalte gestaltet sich damit über die Sprachen hinweg sehr ähnlich.



Grafik 63a und 63b: Schülerschaft - Inhalte der einzelnen Prüfungsteil

Die Abläufe in Englisch werden mit 3,5 (Warm-up), 3,18 (Monolog) und 3,59 (Dialog) bewertet. Auch in Französisch wird die Prise de contact mit 3,54 von 5 im Mittel bewertet, der Monolog mit 3,38 von 5 und der Dialog mit 3,61 von 5. Die Schülerinnen und Schüler in beiden Sprachen bewerten also Ablauf und Inhalte überwiegend *neutral/angemessen* mit einer deutlichen Tendenz zu *gut/eher positiv*.

In beiden Teilen sind extrem schlechte Bewertungen die Ausnahme (mit jeweils ca. 1-2%)



Grafik 64a und 64b: Schülerschaft - Ablauf des Monologs / Dialogs

Die Transparenz der Note ist nach Ansicht der Schülerinnen und Schüler verbesserungsfähig; sie wird von ihnen meist im Mittelfeld (Französisch 2,92; Englisch 2,95) eingestuft. Dennoch halten im Ergebnis 58% der Französisch- und 59% der Englischschülerinnen und -schüler die Note für vergleichbar. Wo dies nicht der Fall ist, wird als Grund überwiegend die Partnerzuordnung genannt; Erfolg sei „von Auslosung oder Harmonie mit dem Partner abhängig.“ (F) Hier einige weitere Einzelstimmen:

- „Partner beeinflusst die Note stark“ (F)
- „Wenn man Partner kennt & mit ihm befreundet ist, lockerer Umgang miteinander & Aufregung geringer“ (F)
- „Dialog hängt vom Partner ab (schlechte Englischkenntnisse oder egoistische Haltung)“ (E)

In diesem Kontext wird auch explizit das Losverfahren thematisiert, und zwar positiv wie negativ:

- „Erfolg ist von dem Harmonieren der Partner abhängig. Das Lösen ist unangebracht und sinnlos“ (E)
- „Unterschiedliche Schwierigkeitsgrade, Lösen faire Auswahl der Partner“ (E)
- „Wo besteht das Problem sich seinen Wunschpartner auszusuchen?“ (E)

Insbesondere bei unterschiedlichem Niveau der Partner gibt es Kritik:

- „Bei Kombination schlechter und guter Partner können Kommunikationsprobleme auftreten, die sich auf schlechteren Partner negativ auswirken können.“ (E)

Weiterhin wird die Reihenfolge der Prüflinge innerhalb der Prüfung (A und B) genannt, wodurch die Prüflinge unterschiedlich viel Vorbereitungszeit erhalten:

- „Leistung abhängig vom Prüfungspartner; Reihenfolge des Monologteils (Partner B hat mehr Zeit zum Nachdenken als Partner A)“ (E)

Auch die Tageszeit der Prüfung wirkt sich nach Ansicht der Schülerinnen und Schüler merklich auf die Vergleichbarkeit der Note aus:

- „Prüfung am Nachmittag, Prüfer nicht mehr aufnahmefähig, Prüflinge nervös, unfair für Note“ (E)
- „Reihenfolge! Manche hatten gar keine Zeit sich vorzubereiten und andere hatten den ganzen Nachmittag Zeit, bei der Benotung wurde dies nicht berücksichtigt.“ (E)

Der Charakter der Prüflinge selbst wird ebenfalls ins Gespräch gebracht; so merkt ein Schüler an, dass „Vielredner im Vorteil“ (E) seien. Auch die unterschiedliche Schwierigkeit der Themen wird erneut thematisiert: „Bilder, Texte, Karikaturen unterscheiden sich stark in Schwierigkeit“ (E) oder „Themen wie "Climate Change" sollten vermieden werden; eher aktuelle Themen.“ (E)

Dreierprüfungen werden überwiegend negativ gesehen:

- „Größere Gruppen machen gleichen Redeanteil schwieriger.“ (E)
- „3er Prüfung schwerer als 2er, weil man länger konzentriert bleiben muss.“ (E)

Zur Dreierprüfung gibt es allerdings auch einzelne positive Stimmen, etwa: „3er Prüfung leichter, näher an realistischer Diskussion, Redefluss erliegt nicht.“ (E)

Im Vergleich zur letzten Zeugnisnote im Fach Französisch geben 23% der Befragten keine Änderung an; in 30% der Fälle war die Note 1-3 Punkte besser, in 6,7% sogar 4-6 Punkte. In 18% der Prüfungen war die Note 1-3 Punkte schlechter, in 8,6% 4-6 Punkte schlechter. Im Fach Englisch sind die Abweichungen ausgeprägter, aber meistens im positiven Bereich: 18% der Befragten geben keine Änderung an, während in 49% der Fälle die Note 1-3 Punkte abwich (bei 30% nach oben und 19% nach unten), in 12% 4-6 Punkte (bei 7% nach oben und 5% nach unten). Lediglich in 2,5% der Fälle ergab sich eine Abweichung um mehr als 6 Punkte (1,3% nach oben, 1,2% nach unten). 17,6% konnten allerdings im Fach Englisch noch keine Angaben machen.

Als Gründe für schlechteres Abschneiden werden hauptsächlich Nervosität, aber auch fehlende Sprachkompetenz genannt, als Gründe für besseres Abschneiden eine höhere mündliche Kompetenz. Trotzdem muss hier berücksichtigt werden, dass die Sprechprüfung eine spezifische

Teilkompetenz abprüft und daher nur unter Vorbehalt mit einer allgemeineren Zeugnisnote, bestehend aus verschiedenen Teilkompetenzen, verglichen werden kann.

3.3.4 Fazit und Anregungen der Schülerinnen und Schüler

Der am häufigsten von den Schülerinnen und Schülern in Fazit geäußerte Wunsch ist eine längere Vorbereitungszeit vor der Prüfung (insbesondere vor dem dialogischen Teil), oft verbunden mit kleineren organisatorischen Verbesserungen:

- „Vorbereitungszeit von etwa 1-2 min. vor Dialog; Handzeichen bei Untergrenze des Zeitlimits, nicht bei Ende“ (E)
- „Bitte bessere Information, mehr Zeit zum Vorbereiten, Dreiergruppen lassen immer einen nicht zu Wort kommen“ (E)
- „Mindestens 5 min. Vorbereitungszeit, um Partner kennenzulernen; ansonsten gut organisiert & geplant“ (E)
- „längere Vorbereitungszeit um sich auf Sprache und Inhalte zu konzentrieren“ (F)

Auch wird angeregt, darauf zu achten, dass die Themen den gleichen Schwierigkeitsgrad haben und aus dem Unterricht bekannt sind.

- „Schwierigkeitsgrad der Prüfungsthemen variiert stark, angleichen“ (E)
- „vergleichbare Themen; bessere Organisation; Bilder bzw. Karikaturen sind meiner Meinung nach angemessen zu erfüllen, während Texte etc. um einiges schwerer sind.“ (E)
- „Manche Themen im Unterricht behandelt, andere nicht; sehr unterschiedliche Schwierigkeitsgrade“ (F)

Zur Verbesserung der Transparenz der Note wird vorgeschlagen, eine ausführliche Bewertung zu erhalten:

- „Hätte gern ein Feedback über meine Leistungen, was gut/schlecht war, um Note nachvollziehen zu können.“ (E)

Wo keine Kritik am Losverfahren an sich geäußert wird, wird eine frühere Bekanntgabe der Partner und des Prüfungstermins angeregt:

- „Man sollte am Vortag erfahren, wann und mit wem man seine Prüfung hat“ (E)
- „Noten transparenter, mehr Übung, Notenbildung nicht unabhängig, sollte nicht im Abitur sein“ (F)
- „Partner vorher sagen, dadurch bessere Vorbereitung“ (F)

Obwohl die deutliche Mehrzahl der Schülerinnen und Schüler das Auswahlverfahren völlig unproblematisch fand, wird wiederholt auch eine Modifikation der Partnerwahl gewünscht. Zwei Alternativen werden genannt: Schülerinnen und Schüler sollten sich die Partner selbst aussuchen können bzw. die Lehrkräfte sollten Partner mit gleichem Niveau zuordnen:

- „Eigene Partnerwahl, dadurch fühlt sich niemand ungerecht behandelt“ (E)
- „Nicht lösen! Lehrer sollte entscheiden, wer auf gleichem Niveau ist“ (E)
- „Lossystem unvoreteilhaft (unterschiedliche Sprachniveaus)“ (F)

Viele Schülerinnen und Schüler loben explizit die gute Organisation:

- „Alles perfekt gelungen“ (F)
- „Die Organisation war gut; ebenso wie die Durchführung. Die Sprechprüfung an sich kann man fairer und besser gestalten“ (E)
- „Organisation war gut, Ablauf strukturiert, also insgesamt positiv“ (E)
- „Organisation, Vorbereitung und Durchführung waren nach meiner Meinung sehr gut.“ (E)
- „Keine; war zufrieden“ (F)
- „langfristige Vorbereitung; Sprechen ist das A&O“ (F)

Andere sind kritischer: „Bitte bessere Information, mehr Zeit zum Vorbereiten“ (E). Negative Stimmen sind insbesondere im Kontext des Zeitdrucks und der oft langen Wartezeit vor der Prüfung am Prüfungstag zu vernehmen.

Schülerinnen und Schüler wünschen sich außerdem, dass Lehrende mehr in die Prüfung eingreifen können: „Lehrer sollten nicht nur zuhören, sondern sich auch beteiligen“(E) – „Lehrer sollte bei Themenverfehlung eingreifen dürfen.“ (E)

Viele Schülerinnen und Schüler befürworteten in ihrem Fazit die Prüfungsform eindeutig:

- „Mehr Vorbereitungszeit, gute Idee mündliche Leistung zu bewerten, sollte schon in Mittelstufe eingeführt werden.“ (E)
- „Wichtiger Bestandteil des Abiturs.“ (E)
- „Idee ist super! Man kann Eng/Frz nur lernen, wenn man es auch spricht.“ (E)
- „Befürchtungen haben sich nicht bewahrheitet, worüber ich sehr froh bin und ein großes Lob aussprechen möchte. Die Einführung der Prüfung kann ich sehr empfehlen.“ (E)
- „Schon ab 5.Kl üben, um besser vorbereitet zu sein“ (F)

Einige wenige Schülerinnen und Schüler warnen vor der Aufnahme der Prüfungen in der vorliegenden Form als Teil des Abiturs. So wird beispielsweise hingewiesen auf die „tolle Alternative zur Kursarbeit, aber in der Form nicht für[s] Abitur.“(E). In einzelnen Freitextfeldern werden drastischere Formulierungen gewählt: „Bitte ersparen Sie die Prüfung den Abiturienten.“ (E) oder: „Völlig unrealistisch. Diese Prüfung hat uns 8 Wochen in der Abivorbereitung gekostet [...].“ (E).

3.4 Konstituierung der Prüfungspaare

Im Folgenden soll ein kurzer Blick auf die Schulen geworfen werden, die nicht am Losverfahren teilgenommen haben, auch wenn die wesentlich schmalere Datenbasis hier keine fundierten Aussagen zulässt.

Zwei Schulen nahmen nicht am Losverfahren teil. Die Fragebögen der Lehrkräfte und Schülerinnen und Schüler dieser Schulen wurden separat ausgewertet und gehen nicht in die Gesamtanalyse ein. Ausgewertet wurden die Antworten von 138 Schülerinnen und Schülern und 18 Lehrkräften. Eine separate Befragung der Schulleitungen der Schulen ohne Losverfahren wurde nicht vorgenommen, da in diesem Fall die Anonymität der Ergebnisse nicht gewährleistet werden konnte. Da die Fragensauswertung sich hier nur auf sprachunabhängige Aspekte bezieht, wurden beide Sprachen gemeinsam berücksichtigt.

Akteure ohne Losverfahren	Englisch	Französisch
Lehrkräfte	16	2
Schülerinnen und Schüler	119	19
Schulleitungen	0	

Hinsichtlich vieler Aspekte sind die Rückmeldungen relativ ähnlich wie bei den Schulen mit Losverfahren. Übereinstimmungen bei den Schülerinnen und Schülern finden sich z.B. beim Vergleich zur letzten Zeugnisnote, bei der Prüfungsvorbereitung und dem Prüfungsformat.

Der Mittelwert der Antwort auf die Frage nach dem Ablauf der Sprechprüfungen insgesamt liegt bei den Schulen mit Losverfahren mit 3,8 von 5 sogar etwas höher als bei den Schulen ohne Losverfahren, allerdings ist hier der Unterschied nicht signifikant. Die Frage nach dem Zurechtkommen mit dem Partner wird in dieser Gruppe anders beantwortet als von den Schülern, deren Prüfungspartner per Los bestimmt wurde. Die folgende Grafik zeigt die Ergebnisse für Englisch und Französisch zusammen (138 Schüler).



Grafik 65: Schülerschaft ohne Losverfahren - Zurechtkommen mit dem Partner

Diese positive Einschätzung ist nach Ansicht der Schülerinnen und Schüler vor allem auf zwei Faktoren zurückzuführen, die unabhängig von der Prüfungsform an sich gesehen werden: Hier werden erstens das Vertrauensverhältnis mit dem jeweiligen Prüfungspartner genannt und zweitens die Möglichkeit, vorher mit den Partnern zu üben. Einige Einzelstimmen dazu bringen dies zum Ausdruck:

- „Sehr positiv, dass man sich Partner aussuchen konnte, minderte Aufregung, man konnte sich besser vorbereiten.“ (E)
- „Langjähriger Freund, Prüfungsangst und Nervosität wirkten dadurch weniger schlimm, angenehmes Klima.“ (E)
- „Konnten uns gut absprechen, deswegen bessere Prüfung.“ (F)

Aber auch hier sind nicht alle Stimmen ausschließlich positiv und verdeutlichen, dass auch ein vertrautes Verhältnis zwischen den Partnern nicht unbedingt eine lebendigere Diskussion bei der Sprechprüfung garantiert: „Freundin, nicht sehr gesprächig.“ (F) oder „Im Dialogteil hat sie öfters gestockt, lange gebraucht um etwas zu sagen, wodurch ich weniger gesagt habe“ (E)

Bei den befragten Lehrkräften fällt die Bewertung der Einführung der Sprechprüfung insgesamt leicht positiver aus als bei den Schulen mit Losverfahren.

Insbesondere bei der Bewertung des Losverfahrens fällt ein hoher Anteil von nicht beantworteten Fragen auf. Unter den Einzelstimmen werden häufig die besseren Übungsmöglichkeiten im Vorfeld hervorgehoben:

- „Ohne Losverfahren haben SuS Möglichkeit gezielt mit Partner zu üben.“ (E)
- „Verstehe nicht, warum sich SuS nicht gemeinsam vorbereiten sollen.“ (E)

Die übrigen Antworten weichen nicht signifikant von denen der Lehrkräfte an Schulen, die das Losverfahren angewendet haben, ab.

Insgesamt sind die aufgezeigten Unterschiede im Meinungsbild wegen der viel geringeren Datenbasis nur äußerst bedingt interpretierbar. Die Zusammensetzung der Prüfungspaare kann nicht als Einzelaspekt isoliert betrachtet werden, sondern muss im Faktorengefüge der anderen Teilaspekte der Sprechprüfung gesehen werden.

4 Zusammenfassung der Ergebnisse und Empfehlungen

4.1 Wahrnehmung durch die Akteure

Die Ergebnisse der einzelnen Auswertungen über alle Akteursgruppen hinweg lassen den Schluss zu, dass die Erprobung der landeszentralen Sprechprüfung insgesamt erfolgreich verlaufen ist. Alles in allem ist eine positive Haltung der Mehrheit aller beteiligten Akteure (Schulleitungen, Lehrkräfte, Schülerinnen und Schüler) erkennbar. Die Stärkung der mündlichen Komponente wird begrüßt; sie wird als wichtige Teilkompetenz mit positiven Effekten für den Fremdsprachenunterricht verbunden. Lehrkräfte und Schulleitungen erkennen an, dass die Durchführung der Prüfungen die mündliche Kompetenz stärkt und heben mehrheitlich die grundlegende Bedeutung der Sprechprüfung für das Fremdsprachenlernen hervor. Seitens der Schülerinnen und Schüler ist nach Lehrerangaben vor der Prüfung eine gewisse Skepsis festzustellen; die Haltung nach der Prüfung ist aber überwiegend positiv.

Überlagert werden diese grundsätzlich positiven Einschätzungen durch Bedenken und Kritikpunkte vorrangig im organisatorischen Bereich, die insbesondere von Schulleitungen und Lehrkräften angesprochen werden. So zeigt sich bei der konkreten Frage nach der Einführung der Prüfung im vorliegenden Format ein ambivalentes Bild, und auch diejenigen Schulleitungen und Lehrkräfte, die die Prüfung an sich befürworten, machen Optimierungsbedarf hinsichtlich der Organisation und des Ablaufs der landeszentralen Sprechprüfung geltend. Auch Schülerinnen und Schüler äußern diesbezüglich Kritik.

Betrachtet man die Ergebnisse differenziert nach Sprachen, fallen neben einigen Konvergenzen (z.B. hinsichtlich der Einschätzung des Schwierigkeitsgrades der einzelnen Prüfungsteile; vgl. Grafik 63a und 63b, 64a und 64b) auch Divergenzen auf. Die Einschätzungen und Bewertungen der Französischlehrkräfte fallen in einigen Fällen im Vergleich etwas positiver aus als die der Englischlehrkräfte, etwa hinsichtlich des Aufwands bei der Vorbereitung im Vergleich zur mündlichen Kursarbeit (Grafik 12a und 12b), der Schülerreaktion auf die Ankündigung der Sprechprüfungen (Grafik 16a und 16b), der Auswirkung auf die Motivation (Grafik 17a und 17b), des Prüfungserlebens (Grafik 18a und 18b), des Bewertungsrasters (Grafik 34a und 34b) und schließlich der möglichen Einführung der Sprechprüfungen ins Abitur (Grafik 40a und 40b). Mehrfach ist festzustellen, dass Wahrnehmungen oder Bewertungen seitens verschiedener Akteure übereinstimmen. Dies ist etwa der Fall im Hinblick auf den Schwierigkeitsgrad der Aufgaben, welcher sowohl von den Lehrkräften (Grafik 23a und 23b), wie auch von der Schülerschaft (Grafik 55a und 55b) als überwiegend angemessen empfunden wird, ebenso hinsichtlich der Klarheit der Aufgaben (Grafik 25a und 25b sowie 62a und 62b). Diese Stellungnahmen weisen klar auf insgesamt gelungene Elemente des Probedurchlaufs der landeszentralen Sprechprüfungen hin.

Umgekehrt deutet die übereinstimmende Nennung von Kritikpunkten seitens der verschiedenen Akteursgruppen auf Änderungsbedarf hin, der bei einer Einführung dieses Prüfungsformats in die Abiturprüfung berücksichtigt werden sollte (s.u. 4.2, 4.3).

Die Korrelation von Ergebnissen der verschiedenen Beteiligten kann schließlich auch Erkenntnispotential im Hinblick auf den Umgang mit kontroversen Aspekten bergen. Dies zeigt sich insbesondere hinsichtlich der Konstituierung der Prüfungspaare bzw. -gruppen: Das unter Lehrkräften und Schulleitungen festgestellte kontroverse Meinungsbild zum Losverfahren mit mehrheitlich starken Befürwortern, aber auch deutlich ausgeprägten Kritikern, wird durch das in der Schülerschaft festgestellte überwiegend hervorragende Urteil gewissermaßen relativiert (vgl. Grafik 54a und 54b); die geäußerten Verbesserungswünsche sollten in die Optimierung dieses Verfahrens einfließen (vgl. unten Kap. 4.4).

Auch wenn in der Mehrheit die Sprechprüfungen insgesamt positiv gesehen werden, sind bei einigen Akteuren Vorbehalte zu beobachten, die zwar im Einzelnen quantitativ nicht signifikant sein mögen, aber die Umsetzung einer Reform zweifellos erschweren können.

4.2 Vorbereitung und Organisation

Die inhaltliche Vorbereitung der Prüfungen wird seitens der Schulleitungen und Lehrkräfte ganz überwiegend als erfolgreich eingeschätzt. Beide Gruppen loben die Qualität der zur Verfügung gestellten Materialien, die von der Mehrzahl der Lehrkräfte extensiv zur Schülervorbereitung genutzt wurden. Dennoch wird auch der Wunsch nach weiteren Materialien geäußert. Auch die Schülerinnen und Schüler fühlen sich insgesamt gut vorbereitet.

Kritik wird von den meisten Akteuren vor allem an dem als hoch empfundenen Verwaltungsaufwand und der kurzfristigen Information der Schulen über die konkrete Erprobung geübt. Im Einzelnen besteht der Wunsch nach mehr Vorlaufzeit, früherer und ausführlicherer Information, mehr Fortbildungen, mehr Vorbereitungsmaterialien und ausreichend Zeit für die Lehrkräfte, um sich mit den Prüfungsunterlagen vertraut zu machen. Auch der Wunsch nach mehr Flexibilität bei der Durchführung, v.a. bei Dreierprüfungen und für den Krankheitsfall von Schülern/Lehrern wird geltend gemacht.

Es spricht aus unserer Sicht nichts dagegen, die geäußerten Kritikpunkte zu überdenken und ggf. einzelne Aspekte zu modifizieren, soweit dies mit dem übergeordneten Ziel einer standardisierten Prüfung konform ist. Schließlich resultieren sie aus der eingehenden Beobachtung der Abläufe seitens der mit den Anforderungen, Inhalten und Abläufen am besten vertrauten Akteure. Es sollte etwa in Erwägung gezogen werden, mehr Zeit für die Notenfindung vorzusehen, die Abläufe im Kontext des Protokollierens zu überarbeiten und das Bewertungsraster formal zu vereinfachen und zu vereinheitlichen.

Die Frage der Konstituierung der Prüfungspaare ist ausgesprochen zentral, daher gebührt ihr besondere Aufmerksamkeit. Das vor allem in der Gruppe der Lehrkräfte (und teils auch von den Schulleitungen) ambivalent gesehene Losverfahren sollte eingehend von verschiedenen Seiten beleuchtet und gegebenenfalls angepasst werden. Der Gruppe von etwa 20% bei den Französisch- und 25% bei den Englischlehrkräften, die das Losverfahren insgesamt negativ bewerten, steht eine ganz überwiegend positive Rückmeldung der Schülerinnen und Schüler gegenüber, bei denen es nur in wenigen Fällen Schwierigkeiten mit dem zugelosten Partner gab. Die Schülerinnen und Schüler geben zu mehr als der Hälfte an, sogar *hervorragend* mit ihrem Partner zurechtgekommen zu sein. Letztlich ist jeder Fall eines Prüflings, der angibt, *schlecht* oder *sehr schlecht* mit dem Partner zurechtgekommen zu sein, unglücklich. Unabhängig von der Zusammenstellung der Prüfungspaare spielen hier viele verschiedene, vor allem individuelle Faktoren, eine Rolle. Ein „Totalausfall“ (vgl. Kap.3.2.3.5) lässt sich sicher mit keinem Format völlig vermeiden.

Angesichts der zahlreichen Aspekte, die grundsätzlich für das Losverfahren sprechen⁶, sollte eine Entscheidung darüber daher durch verschiedene Maßnahmen flankiert werden. Wichtig ist es zunächst, die Anforderungen an eine Sprechprüfung und die mit ihr verbundenen spezifischen Teilkompetenzen allen Akteuren transparent zu machen (Unterstreichen des Fokus

⁶ Stichpunktartig seien hier Ergebnisse verschiedene wissenschaftlicher Studien zusammengefasst: Gleiche Bedingungen werden erfahrungsgemäß leichter akzeptiert als entweder Zuteilungen durch Lehrkräfte (z.B. nach Extrovertiertheit/ Introvertiertheit) oder selbst ausgesuchte Partner, weil favorisierte Sprechpartner nicht allen zukommen können (vgl. z.B. Herdegen 2009:209). Potentiell prüfungsrelevante Unterschiede in Charaktereigenschaften von Prüflingen bzw. das Problem unterschiedlicher Kompetenzniveaus können darüber hinaus auch bei selbst gewählten Partnern zum Tragen kommen (Davis 2009). Entsprechende Studien (z.B. Gan 2011) finden keinen statistisch relevanten Zusammenhang zwischen Unterschieden zwischen Charaktereigenschaften in Prüfungspaaren und der Bewertung. Ein weiteres Argument für das Losverfahren liegt im Prüfungsziel. Dieses besteht darin, Rückmeldung dazu zu geben, wie der Prüfling in für ihn relevanten authentischen kommunikativen Situationen (die sich durch eine je eigene, nicht vorhersehbare Dynamik charakterisieren; vgl. Brooks 2009) möglichst erfolgreich agiert. Daher sollten die Prüfungsbedingungen sich einer solchen Situation maximal annähern, um Rückschlüsse auf die geprüfte Kompetenz ziehen zu können (Weir 2004). Schließlich fördert das Zufallsverfahren indirekt durch die Notwendigkeit der Vorbereitung auf mehrere potentielle Partner Kooperation im Unterricht und verhindert die Konzentration auf wenige feste Partner (vgl. Saville & Hargreaves 1999).

auf mündliches Interagieren vs. vertieftes Auseinandersetzen mit Inhalten), etwa im Rahmen von Fortbildungen. Dreh- und Angelpunkt einer gelungenen Sprechprüfung ist schließlich eine gut geschulte und vorbereitete prüfende Lehrkraft.

Die Schülerinnen und Schüler sollten den kommunikativen Umgang mit verschiedenen, nicht selbst gewählten Gesprächspartnern (sowohl mit etwa gleichem als auch mit unterschiedlichem Sprachniveau) verstärkt im Unterricht trainieren. Hierbei sollte der Annahme entgegengewirkt werden, dass ein gleiches Niveau prinzipiell mit Fairness korreliert werden kann, wie es einige der Akteure andeuten: Die Sprechprüfung zielt nicht auf die Abfrage von konkretem inhaltlichem Wissen oder einer exhaustiv ausgearbeiteten Interpretation der zur Verfügung gestellten Materialien ab, sondern vielmehr darauf, die mündliche Kompetenz der Schülerinnen und Schüler in einer authentischen Sprechsituation zu überprüfen. Hier kann daher verstärkt darauf hingewiesen werden, dass Strategien zur Aufrechterhaltung des Gesprächsflusses und das aktive Einbeziehen des Partners in die eigenen Überlegungen ebenso positiv bewertet werden können, wie die konkrete Nennung von Interpretationselementen. Bei einer Paarung mit unterschiedlichen Niveaus kann dabei vor allem der/die stärkere Schüler/in dazu angeregt werden, die Gesprächsleitung in die Hand zu nehmen und durch gezieltes Nachfragen oder Verhandeln des Diskussionsverlaufs sowohl seine/ihre eigenen Kompetenzen unter Beweis zu stellen, als auch den/die schwächere/n Partner/in zur aktiveren oder tiefer gehenden Gesprächsteilnahme anzuregen.

Die geäußerte Kritik kann hier, bei Beibehaltung einer zentralisierten Prüfung, konstruktiv in die Überlegungen zur Optimierung der Zusammenstellung der Prüfungspaare einbezogen werden, z.B. die frühere Bekanntgabe der zugelosten Partner und des genauen Prüfungszeitpunkts.

4.3 Prüfungsformat

Das Format der Prüfungen wird insgesamt als zweckdienlich bewertet. Die Prüfungsmaterialien werden überwiegend als vergleichbar angesehen, wenn auch der Schwierigkeitsgrad insbesondere der Tabellen- und Karikaturimpulse von den Lehrkräften als höher eingeschätzt wird als der der Text- oder Bildimpulse. Diese Sicht deckt sich im Wesentlichen mit der der Schülerinnen und Schüler, von denen allerdings Karikaturen im Vergleich zu Tabellen als insgesamt leichter bewertet werden. In zukünftigen Prüfungen sollte darauf geachtet werden, die Kluft zwischen Text und Bild einerseits und Tabelle und Karikatur andererseits zu verringern. Wie sich im hier untersuchten Probedurchlauf gezeigt hat, ist die gezielte Arbeit mit den verschiedenen Impulsen anhand der zur Verfügung gestellten Prüfungsmaterialien ein Erfolg versprechender Weg, um die Prüflinge mit den kommunikativen Anforderungen der einzelnen Typen vertraut zu machen und Routinen entstehen zu lassen (auch wenn Prüfungen und Prüfungsmaterial immer eine nicht vermeidbare individuelle Komponente enthalten). Dabei sollte nicht aus den Augen verloren werden, dass sprachliches Handeln generell das Verfügen über sprachliche Mittel erfordert. Unabhängig von der konkreten Prüfungsvorbereitung (vgl. Kap. 3.2.1. und 3.3.1) ist das Einüben der sprachlichen Mittel empfehlenswert. Diese sollten überdies dem aktuellen Sprachgebrauch entsprechen.

Die Kritik an der in Einzelfällen mangelnden Passung von Themenbereichen der Prüfungen und Lehrplan erscheint berechtigt. Es ist jedoch davon auszugehen, dass für weitere Durchläufe die Themenbereiche der Prüfungen durchgehend vom Lehrplan abgedeckt werden.

Die Vereinfachung der Bewertungsraster, deren weniger komplexe Formulierung, trennschärfere Kategorien und eine bessere Anpassung an die verschiedenen Sprechimpulse sollten zu einer Verbesserung der Abläufe beitragen. Die wiederholt angesprochene Bewertungsproblematik im Bereich *Interaktion* (vor allem im Rahmen des monologischen Prüfungsteils) wird der Prüfung inhärent bleiben, kann aber durch eine entsprechende Überarbeitung der Vorgaben entschärft werden. Auch ein einheitliches Bewertungsraster für Englisch und Französisch erscheint sinnvoll.

In diesem Kontext sollte im Sinne einer erhöhten Transparenz auch die Erstellung eines abschließenden gemeinsamen Protokolls und die Möglichkeit zur Einsichtnahme für die geprüften Schülerinnen und Schüler überdacht werden.

Schließlich sollte auch die Frage des Notenschemas noch einmal gründlich abgewogen werden, dessen Überarbeitung (etwa bezogen auf eine 15 Punkte-Skala, einzelne Kategorien zum Ankreuzen) ebenfalls häufiger angeregt wird.

Der Kritik an dem als hoch empfundenen bürokratischen Aufwand könnte durch die insbesondere von vielen Lehrkräften angeregte Modifikation der Protokollführung (z.B. ein gemeinsames Protokoll zur Entlastung der/des Erstprüfenden, ggf. sogar ein zusätzlicher Protokollant) begegnet werden. Hinsichtlich der zudem von einigen Lehrenden geforderten weniger starren Vorgaben des Prüfungsskriptes (mehr Flexibilität bei Lehrerimpulsen) ist eine Abwägung zwischen den Erfordernissen des standardisierten Prüfungsformats und dem individuellen Handlungsspielraum im Hinblick auf das Gewährleisten eines erfolgreichen Prüfungsablaufs vorzunehmen. Auch hier kommt der Prüferschulung eine zentrale Rolle zu.

4.4 Durchführung

Schülerschaft, Lehrkräfte und Schulleitungen bewerten den konkreten Prüfungsablauf überwiegend positiv. Alle drei Gruppen formulieren allerdings Nachbesserungsbedarf aus ihrer jeweiligen Sicht.

Seitens der Schulleitungen wird vor allem das Nachsteuern bei der Prüfungsorganisation angesprochen. Im Hinblick auf die wesentliche Rolle der (erweiterten) Schulleitung bei der Ein- und Durchführung der landeszentralen Sprechprüfungen ins Abitur sollten die vorgeschlagenen Modifikationen (vgl. Kap. 3.1.4) eingehend geprüft werden.

Für die Lehrkräfte stehen besonders Fragen nach der zeitlichen Strukturierung der Prüfungstage im Vordergrund. Kritisiert wird der Zeitdruck für alle Beteiligten, bedingt durch die enge Prüfungstaktung und die hohe Prüfungszahl pro Tag; gefordert werden hier vor allem mehr Zeit für die Notenfindung und Pausen zwischen den einzelnen Prüfungen, aber auch mehr Zeit für die Vorbereitung nicht nur für die Prüflinge, sondern auch für die Lehrenden.

Teilweise ergeben sich hier Übereinstimmungen mit den Schülerinnen und Schülern, die vielfach eine frühere Bekanntgabe der Prüfungspaare und – unter dem Stichwort *Zeitdruck* (vgl. z.B. Kap. 3.3.4) – mehr Vorbereitungszeit während der Prüfung anregen. Zudem wird die Erläuterung und zeitnahe Kommunikation der Note gewünscht, was sich mit der angesprochenen Anregung der Lehrkräfte, ein gemeinsames Protokoll zur Einsichtnahme einzuführen (vgl. Kap. 4.3), deckt.

Sicherlich sind – wie bei anderen schulischen Prüfungsformaten auch – immer Abwägungen zwischen als ideal und wünschenswert angesehenen Umständen und dem organisatorisch Machbaren (z.B. mehr Zeit einplanen) zu treffen; die geäußerten Verbesserungsvorschläge stehen aber grundsätzlich nicht dem Ziel der landeszentralen Sprechprüfungen entgegen, sondern können zur Optimierung des Verfahrens beitragen. Ihre mögliche Umsetzung sollte daher eingehend geprüft werden. Insgesamt sollte angesichts der großen Belastung der Schulleitungen wie der Lehrkräfte durch organisatorisch-administrative Aufgaben ein Weg gefunden werden, um einerseits mit den Sprechprüfungen die Basis der fremdsprachlichen Kompetenzen der Schülerinnen und Schüler im Saarland zu verbreitern, andererseits die zusätzliche Belastung möglichst gering zu halten⁷.

⁷ So erscheint es etwa sinnvoll, bei entsprechender Prüfungszahl weiterhin zwei Tage für die verschiedenen Teilprüfungen beizubehalten und sie ggf. nach den schriftlichen Abiturprüfungen anzusetzen, um den allgemeinen Unterrichtsablauf weniger zu beeinträchtigen.

5 Schlussfolgerungen und Ausblick

Die Ergebnisse zeigen, dass alle Akteure einer Einführung der Sprechprüfungen trotz skeptischer Einzelstimmen und punktuell formulierten Nachbesserungsbedarfs generell positiv gegenüberstehen. Die Kritikpunkte zeigen eine Bandbreite an als problematisch identifizierten Aspekten auf. Die von den verschiedenen Akteuren geäußerten Anregungen können insbesondere dort auf ihre Angemessenheit geprüft und ggf. weiter elaboriert und ergänzt werden, wo sie mit dem gemeinsamen Ziel einer Einführung zentraler, standardisierter Sprechprüfungen übereinstimmen. In einigen Bereichen können darauf basierend Maßnahmen zur Optimierung des Sprechprüfungsverfahrens ausgearbeitet werden.

Verschiedene Antworten belegen, teils implizit, den mit diesem Prüfungsformat verbundenen Wert für den Fremdsprachenunterricht, etwa wenn geäußert wird, dass mit den Sprechprüfungen spezifische Teilaspekte bewertet werden, die schriftlich nicht bewertbar sind (vgl. Kap. 3.2.3.1). Sicher müssen die Schülerinnen und Schüler stärker mit dem spezifischen Format vertraut gemacht werden; dabei sollte im Unterricht der Akzent verstärkt auch auf Kommunikation unter Schülerinnen und Schülern statt ausschließlich/überwiegend zwischen Lehrkraft und Schüler gelegt werden. Wenn Lerner regelmäßig ermutigt werden, mit unterschiedlichen Partnern kommunikativ zu interagieren, erscheint auch das unter den Lehrkräften so kontrovers beurteilte, von den Schülern aber als weitgehend unproblematisch empfundene Losverfahren in einem anderen Licht. Im Unterricht regelmäßig praktiziert – wie einige Lehrkräfte dies im Rahmen der Erprobung der Sprechprüfungen mit Erfolg getan haben –, ist es ein Mittel, um die Bandbreite unterschiedlicher Kommunikationspartner und im Klassenraum weitgehend authentische Kommunikationsbedingungen in der Fremdsprache kennenzulernen. Auch für eine gleichmäßigere Kompetenzentwicklung innerhalb der Klasse dürfte sich dies als vorteilhaft erweisen.

Es ist zweifellos eine Herausforderung, die Komplexität mündlicher Kommunikationssituationen für zwei Fremdsprachen gleichzeitig in einem standardisierten Prüfungsformat abzubilden und in den schulischen Alltag zu integrieren. Aus einer Jahrhunderte langen Tradition heraus ist der Stellenwert der Schriftlichkeit im Fremdsprachenunterricht auch heute noch relativ hoch, auch wenn mit der Veröffentlichung des *Gemeinsamen Europäischen Referenzrahmens für Sprachen* (GER) im Jahre 2001 einige Veränderungsprozesse in Gang gekommen sind. Der Stellenwert von Mündlichkeit im Fremdsprachenunterricht hat bedeutend zugenommen. Die in der vorliegenden Studie dokumentierten bereits bestehenden Erfahrungen mit Mündlichkeit im Englisch- und Französischunterricht (inkl. ihrer Überprüfung, z.B. in Form von mündlichen Klassenarbeiten) sind bereits ein Ergebnis dieses Wandlungsprozesses, ebenso die mehrfach erwähnten Hörverstehensprüfungen. Sie verdeutlichen die zeitliche Dimension solcher Prozesse. Vieles, was bei der Einführung des GER kritisch wahrgenommen wurde, ist mittlerweile gut im saarländischen Fremdsprachenunterricht verankert, beispielsweise die Kompetenzorientierung und -differenzierung, das Anstreben einer funktionalen Kompetenz als Zielnorm oder das sprachenvernetzende Unterrichten. Die landeszentralen Sprechprüfungen in Französisch und Englisch tragen diesen bildungspolitischen Leitlinien gleich mehrfach in innovativer Weise Rechnung: durch den Fokus auf die Sprech- und Kommunikationskompetenz, aber auch dadurch, dass die beiden modernen Fremdsprachen fächerübergreifend nach derselben Maßgabe behandelt werden, so dass gemeinsame Standards für Englisch und Französisch entstehen.

Die bereits im Kontext des GER genannten Prinzipien sind im schulischen Fremdsprachenunterricht, ebenso wie im Sprachenkonzept des Saarlandes (*Neue Wege zur Mehrsprachigkeit im Bildungswesen*, 2011), präsent und fließen in sprachliche Lehr- und Lernprozesse auf allen Bildungsstufen ein. Die damit angestoßene Neujustierung des Fremdsprachenunterrichts entspricht einem komplexen Wechselspiel, in dem viele Faktoren auf unterschiedlichen Ebenen zusammenwirken. Der Fremdsprachenunterricht ändert sich dabei durch verschiedene Einflussmomente. Einerseits sind im Zuge der Kompetenzorientierung neue Teilkompetenzen relevant geworden, die zu neuen Prüfungsformaten führen („bottom up“);

andererseits können Veränderungen auch durch das Einführen neuer Prüfungsformate hervorgerufen werden („top down“)⁸. Die Ergebnisse der vorliegenden Studie lassen an verschiedenen Stellen solche Umstellungen erkennen (z.B. wenn Schülerinnen und Schüler systematisch mit den verschiedenen Impulsen vertraut gemacht werden). Die landeszentralen Sprechprüfungen erscheinen somit einerseits als logische Konsequenz, andererseits als Motor des beschriebenen Wandels. Wichtig ist ein solcher Motor nicht zuletzt im Kontext der Frankreichstrategie des Saarlandes, deren Ziel ja „[...] ein leistungsfähiger mehrsprachiger Raum deutsch-französischer Prägung“ (Eckpunktepapier 2013:4) ist.

Profunde Kommunikationsfähigkeit und kommunikative Kompetenzen sowohl in der Nachbar- und Partnersprache Französisch als auch in der globalen *lingua franca* Englisch sind unverzichtbar, um den saarländischen Schülerinnen und Schülern optimale Voraussetzungen für einen guten Einstieg in die Arbeitswelt mit auf den Weg zu geben. Die Sprechprüfung spielt hier eine wichtige Rolle, um Schülerinnen und Schüler auf die Auseinandersetzung mit authentischen Kommunikationssituationen in Alltags- und Arbeitskontexten vorzubereiten. Aus diesem Grund erscheint uns das bisweilen geäußerte Szenario, mit der Einführung der Sprechprüfungen in das Abitur die dort bislang neben der Schreibaufgabe vorhandene Hörverstehensprüfung zu ersetzen, kein Schritt in die richtige Richtung. Für die Erlangung einer soliden funktionalen kommunikativen Kompetenz ist das gesamte Spektrum an Kompetenzen nötig (Hören, Sprechen, Lesen, Schreiben). Auch einzelne Teilkompetenzen tragen hierbei durch Synergieeffekte zu einer übergreifenden kommunikativen Kompetenz bei. Es sollte also kein Auspielen der verschiedenen Kompetenzbereiche gegeneinander stattfinden.

Die insgesamt positiven Ergebnisse der probeweise durchgeführten Sprechprüfungen ermutigen dazu, den eingeschlagenen Weg weiter zu beschreiten und die in einigen Bereichen festgestellten Probleme als ‚Kinderkrankheiten‘ zu sehen, wie sie stets mit innovativen Prozessen verbunden sind, die aber keine grundlegenden Hindernisse darstellen.

⁸ Dabei ist zu berücksichtigen, dass an ein neues Prüfungsformat und das Abprüfen einer ‚neuen‘ Teilkompetenz Fragen gerichtet werden, denen sich etablierte Formate und ‚traditionelle‘ Kompetenzen nicht mehr stellen müssen.

6 Literaturverzeichnis

Börner, W. (1989). *Didaktik schriftlicher Textproduktion in der Fremdsprache. Textproduktion. Ein interdisziplinärer Forschungsüberblick*. Tübingen: Niemeyer, 348-376.

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing* 26 (3), 341-366.

Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing* 26, 367-396.

Europarat (Hrsg.) (2001). *Gemeinsamer Europäischer Referenzrahmen für Sprachen GER*. Straßburg: Europarat.

Gan, Z. (2011). An Investigation of Personality and L2 Oral Performance. *Journal of Language Teaching & Research* 2 (6), 1259-1267.

Herdegen, P. (2009). *Schulische Prüfungen: Entstehung, Entwicklung, Funktion: Prüfungen am bayerischen Gymnasium vom 18. bis zum 20. Jahrhundert*. Bad Heilbrunn: Julius Klinkhardt.

Ministerium für Bildung Saarland (Hrsg.) (2011). *Sprachenkonzept im Saarland 2011. Neue Wege zur Mehrsprachigkeit im Bildungssystem*. Online:
http://www.saarland.de/dokumente/res_bildung/Das_Sprachkonzept_Saarland_2011.pdf.
(13.01.16).

Saville, N., & Hargreaves, P. (1999). Assessing speaking in the revised FCE. *ELT journal*, 53(1), 42-51.

Staatskanzlei des Saarlandes (Hrsg.). (2014). *Eckpunkte einer Frankreichstrategie für das Saarland*. Saarbrücken: Staatskanzlei. Online:
http://www.saarland.de/dokumente/res_stk/D_Eckpunkte_Frankreich-Strategie_210114.pdf.
(13.01.16).

Weir, C. J. (2004). *Language Testing and Validation. An Evidence-Based Approach*. Basingstoke, New York: Palgrave Macmillan.