

# 3 Prestaties mondelinge taalvaardigheid

**Hoe vaardig zijn leerlingen eind groep 8 op het gebied van luisteren, spreken en gesprekken voeren? In dit hoofdstuk beschrijven we wat leerlingen in groep 8 van het basisonderwijs kennen en kunnen binnen de verschillende subdomeinen van mondelinge taalvaardigheid. Ook besteden we aandacht aan verschillen met de prestaties voor luisteren en spreken in eerder peilingsonderzoek (2007 voor luisteren en 2010 voor spreken).**

## 3.1 Referentiekader en -niveaus

In het Referentiekader Taal en Rekenen is het domein mondelinge taalvaardigheid een apart onderdeel van Taal. Mondelinge taalvaardigheid is opgesplitst in drie subdomeinen: luisteren, spreken en gesprekken voeren (Staatsblad van het Koninkrijk der Nederlanden, 2010). In dit peilingsonderzoek brengen we de vaardigheden van de leerlingen op deze drie subdomeinen in kaart. Daarbij gaan we per subdomein na op welk vaardigheidsniveau leerlingen presteren, de zogenaamde referentieniveaus.

De referentieniveaus beschrijven wat leerlingen, van de basisschool tot aan het hoger onderwijs, moeten kennen en kunnen op de verschillende (sub)domeinen van de Nederlandse taalvaardigheid (Expertgroep Doorlopende Leerlijnen Taal en Rekenen, 2009). De niveaus worden beschreven in twee 'kwaliteiten': de fundamentele kwaliteit (1F) en de streefkwaliteit (2F). Bij de introductie van de referentieniveaus is door de commissie Meijerink (2009) de ambitie uitgesproken dat minimaal 85 procent van de leerlingen het fundamentele niveau beheerst en minimaal 65 procent het streefniveau.

De instrumenten in dit onderzoek meten of leerlingen de referentieniveaus 1F en 2F beheersen, zoals gedefinieerd in het Referentiekader Taal en Rekenen. De beschrijvingen van deze niveaus en een veldraadpleging met domeinexperts, taalcoördinatoren, schoolleiders en leerkrachten, vormden de voornaamste bron bij het samenstellen van de inhoud van de afzonderlijke instrumenten. Voor meer informatie over het instrumentarium, zie hoofdstuk 1 van deel C in dit rapport en Van Langen et al., 2017.)

Om de vaardigheid van leerlingen te kunnen plaatsen in het licht van de referentieniveaus, is daarna in een inhoudelijke standaardbepaling door experts bepaald waar de cesuren of 'grenspunten' liggen voor het beheersen van referentieniveau 1F en 2F (Linthorst, Hemker, Koerhuis, Feskens & Koops, 2018)<sup>15</sup>. Die grenspunten worden in de volgende paragrafen gebruikt om de scores van de leerlingen te interpreteren.

<sup>15</sup> Deze onafhankelijke standaardbepaling is georganiseerd door CvTE en is onder leiding van Stichting Cito uitgevoerd door een expertpanel, bestaande uit achttien personen met verschillende expertise, voornamelijk taalexperts en leerkrachten.

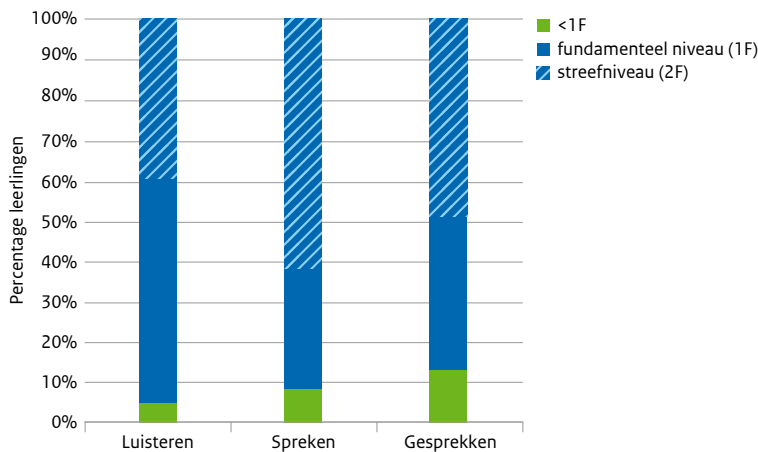
## 3.2 Prestaties

### 3.2.1 De subdomeinen tezamen

In figuur 3.2.1a is weergegeven op welk niveau leerlingen uit groep 8 van het basisonderwijs presteren op de drie subdomeinen van mondelinge taalvaardigheid. Voor spreken en gesprekken geldt daarbij dat de uitspraken over de beheersing van de referentieniveaus geïnterpreteerd moeten worden binnen de context van de afgenomen taken. Ze kunnen niet gegeneraliseerd worden naar andere spreek- of gespreksituaties.

Voor alle subdomeinen geldt dat ten minste 85 procent van de leerlingen het fundamentele niveau 1F beheerst. Respectievelijk 95, 92 en 87 procent beheerst het niveau 1F voor luisteren, spreken en gesprekken. De eerder door de commissie Meijerink gestelde ambitie dat minimaal 85 procent van de leerlingen het fundamentele niveau voor taal en rekenen beheerst, is daarmee voor alle subdomeinen van mondelinge taalvaardigheid behaald. Ook voor lezen, taalverzorging en rekenen werd deze ambitie in groep 8 ruimschoots behaald (Inspectie van het Onderwijs, 2018). Bij geen van de subdomeinen beheerst 65 procent of meer van de leerlingen het streefniveau 2F. De beheersing van het streefniveau voor spreken ligt daar met 62 procent nog het dichtst bij in de buurt. Voor luisteren beheerst 40 procent van de leerlingen het streefniveau. Voor gesprekken is dit 49 procent.

Figuur 3.2.1a – Percentage leerlingen per referentieniveau per subdomein  
( $n_{\text{luisteren}} = 1531$ ,  $n_{\text{spreken}} = 1365$ ,  $n_{\text{gesprekken}} = 1380$ )



### 3.2.2 Luisteren

In het referentiekader is het subdomein luisteren in drie taken uiteengelegd. Dit zijn:

1. *luisteren naar instructies*
2. *luisteren als lid van een live publiek*
3. *luisteren naar radio en televisie en gesproken tekst op internet* (verder in de tekst *luisteren naar radio*)

De taak *luisteren als lid van een live publiek* is in dit onderzoek niet als een afzonderlijke taak meegenomen. Deze vertoont in de (digitale) toetssetting van dit onderzoek namelijk veel overlap met *luisteren naar radio en televisie en gesproken tekst op internet*.

De luistertaken zijn uitgevoerd met behulp van veertien verschillende video- en audiofragmenten. Deze vertonen tekstenkenmerken zoals beschreven in het referentiekader voor niveau 1F en 2F. Zo hebben fragmenten op 1F-niveau bijvoorbeeld een lagere informatiedichtheid dan fragmenten op 2F-niveau. Het onderscheid tussen 1F en 2F komt daarnaast ook tot uiting in de kenmerken van de taakuitvoering; bij sommige fragmenten wordt ingegaan op het *begrijpen* van het fragment, terwijl er bij andere fragmenten een beroep

wordt gedaan op *interpretatie*.<sup>16</sup> Bij *begrijpen* gaat het, afhankelijk van het referentieniveau (1F versus 2F), bijvoorbeeld om het onderscheiden van hoofd- en bijzaken, belangrijke informatie uit de tekst halen, of de betekenis van onbekende woorden afleiden uit de context. Bij *interpreteren* gaat het bijvoorbeeld om relaties leggen tussen tekstuele informatie en eigen ervaringen, het interpreteren van informatie en meningen of het verwoorden van de bedoeling van de spreker (zie verder tabel 3.2.2a).

In totaal kregen leerlingen 30 vragen voorgelegd die horen bij 10 video- en 4 audiofragmenten. 8 vragen gingen over fragmenten waarin instructies werden gegeven. De overige 22 vragen gingen over radio- en televisiefragmenten en gesproken tekst op internet.

Tabel 3.2.2a – Kenmerken van de taakuitvoering luisteren (Staatsblad van het Koninkrijk der Nederlanden, 2010)

Taalkenmerk	Niveau 1F	Niveau 2F
<b>Begrijpen</b>	<ul style="list-style-type: none"> <li>• Kan hoofdzaken uit de tekst halen.</li> <li>• Kan via selectie belangrijke informatie uit de tekst halen en kan de manier van luisteren daar op afstemmen (bijvoorbeeld globaal, precies, selectief/gericht).</li> </ul>	<ul style="list-style-type: none"> <li>• Kan de hoofdgedachte van de tekst weergeven en kan onderscheid tussen hoofd- en bijzaken maken.</li> <li>• Kan relaties tussen tekstdelen leggen.</li> <li>• Kan informatie ordenen (bijvoorbeeld op basis van signaalwoorden) voor een beter begrip.</li> <li>• Kan wanneer nodig de betekenis van onbekende woorden afleiden uit de vorm, woordsoort, samenstelling of context.</li> <li>• Kan beeldspraak herkennen.</li> <li>• Kan een relatie leggen tussen tekst en beeld.</li> </ul>
<b>Interpreteren</b>	<ul style="list-style-type: none"> <li>• Kan informatie en meningen interpreteren voor zover deze dicht bij de leerling staan.</li> <li>• Kan relaties leggen tussen tekstuele informatie en eigen kennis en ervaringen.</li> </ul>	<ul style="list-style-type: none"> <li>• Kan informatie en meningen interpreteren.</li> <li>• Kan de bedoeling van de spreker(s) of het doel van de makers van een programma verwoorden.</li> </ul>

De antwoorden op alle afzonderlijke vragen in de luistertoets konden worden samengenomen tot 1 score per leerling voor het subdomein luisteren, met een maximale score van 30 punten.

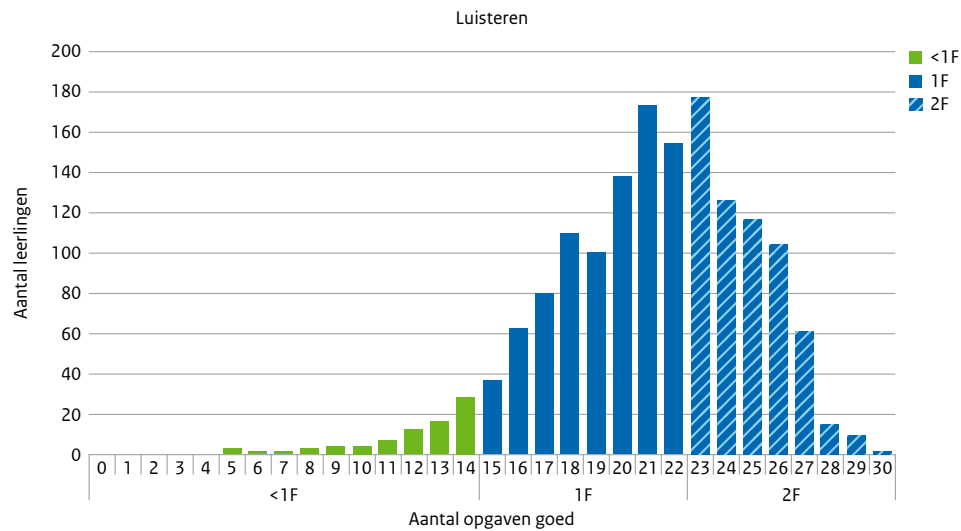
Figuur 3.2.2a laat zien dat er geen leerlingen zijn met een puntenaantal of score onder de 5 en dat 1 leerling de maximale score heeft behaald. Deze leerling heeft dus alle vragen goed beantwoord. De verschillende scores van de leerlingen zijn ingedeeld in drie groepen:

- Leerlingen met een score van 15 of hoger beheersen het fundamentele niveau 1F (blauw in de figuur).
- Leerlingen met een score van 23 of hoger beheersen het streefniveau 2F (blauw met arcering in de figuur).
- Leerlingen met een score lager dan 15 hebben het fundamentele niveau voor luisteren nog niet bereikt (groen in de figuur).

<sup>16</sup> In het referentiekader worden vier taakkenmerken onderscheiden: *begrijpen*, *interpreteren*, *evalueren* en *samenvatten*. In dit onderzoek zijn de kenmerken *evalueren* en *samenvatten* niet getoetst. *Evalueren* is niet getoetst omdat dit alleen maar indirect getoetst kan worden. *Samenvatten* is niet getoetst omdat dit eerder een beroep doet op spreekvaardigheid dan luistervaardigheid (zie ook Van Langen et al, 2017).

1531 leerlingen hebben de luistertoets gemaakt. Ruim 95 procent van de leerlingen beheerst minimaal het fundamentele niveau 1F voor luisteren, 40 procent beheerst ook het streefniveau 2F. De figuur laat ook zien dat voor veel leerlingen het referentieniveau 2F binnen bereik ligt. Voor 154 leerlingen geldt dat zij met één goed antwoord meer op de luistertoets, op referentieniveau 2F zouden presteren. Daarmee zou het percentage leerlingen op niveau 2F stijgen van 40 naar 50 procent.

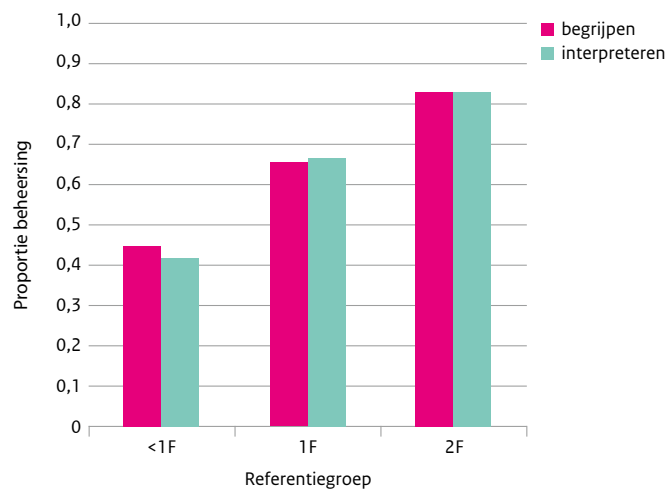
Figuur 3.2.2a – Scores op de toets luisteren (n=1531)



**De afzonderlijke kenmerken voor luisteren**

De twee taken die samen de luistertoets vormen – *luisteren naar instructie* en *luisteren naar radio* - zijn ieder aan de hand van twee kenmerken gemeten: *begrijpen* en *interpreteren*. Als we kijken naar hoe leerlingen hebben gepresteerd op deze twee afzonderlijke kenmerken, zien we geen noemenswaardige verschillen. Ze hebben van zowel de opgaven die ingaan op *begrijpen*, als van de opgaven waarbij ze moeten *interpreteren*, gemiddeld 72 procent goed. We kunnen daarmee stellen dat gemiddeld genomen *begrijpen* als even moeilijk wordt ervaren als *interpreteren*. Als we de prestaties op de twee kenmerken per referentiegroep bekijken, zien we dat er ook binnen iedere referentiegroep nauwelijks verschillen zijn in de beheersing van *begrijpen* en *interpreteren* (zie figuur 3.2.2b).

Figuur 3.2.2b – Gemiddelde proportie beheersing van *begrijpen* en *interpreteren* in referentiegroepen <1F, 1F en 2F (n=1531)



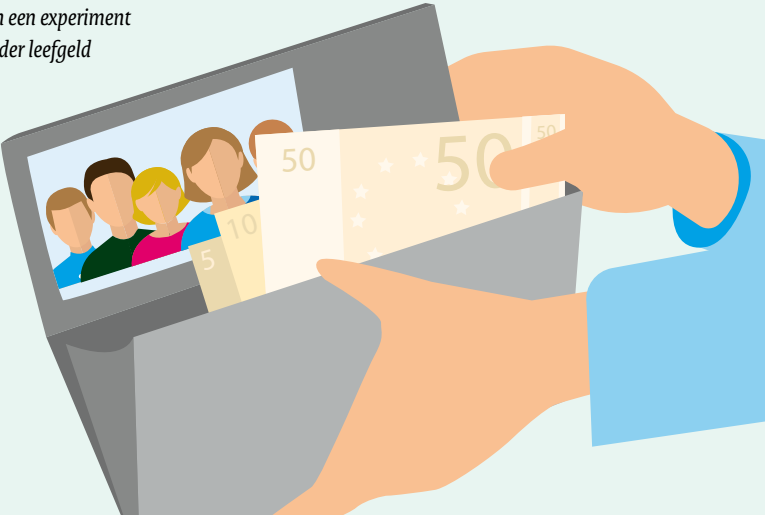
### Wat onderscheidt leerlingen presterend op de verschillende referentieniveaus van elkaar?

Voor iedere referentiegroep (<1F, 1F en 2F) hebben we per opgave in de luistertoets gekeken naar de proportie leerlingen met een goed antwoord, de zogenaamde p'-waarde<sup>17</sup>. Op die manier weten we welke opgaven opvallend vaker goed zijn beantwoord door de ene referentiegroep dan door de andere referentiegroep. Zo hebben we drie opgaven geïdentificeerd die het meest onderscheidend zijn voor referentieniveau <1F versus referentieniveau 1F en drie opgaven die het meest onderscheidend zijn voor referentieniveau 1F versus referentieniveau 2F<sup>18</sup>.

Figuur 3.2.2c geeft een voorbeeld van een opgave waarop de referentiegroepen <1F en 1F grote verschillen in beheersing vertonen.

Figuur 3.2.2c – Fragment - Familie Heuf leeft met heel weinig geld

De familie Heuf doet mee aan een experiment waarin ze een week lang minder leefgeld krijgen (65 euro).



Opgave:	
<b>De familie Heuf krijgt leefgeld. Wat kunnen ze kopen van het leefgeld? Kies de twee goede antwoorden.</b>	
Afwasmiddel	0,5
Bananen	0,5
Een simkaart	0
Een spelletje	0
Sportschoenen	0

In dit fragment moeten leerlingen relaties leggen tussen tekstuele informatie en eigen kennis en ervaringen (interpretatie op niveau 1F). Ze krijgen maximaal 1 punt (0,5 voor afwasmiddel en 0,5 voor bananen). Leerlingen uit de referentiegroep 1F doen dit significant beter dan leerlingen die het fundamenteel niveau voor luisteren nog niet beheersen.

<sup>17</sup> De p'-waarde beschrijft de moeilijkheidsgraad van de vragen in de toets, in termen van de proportie leerlingen die de vraag goed heeft beantwoord. Een p'-waarde van bijvoorbeeld .45 vertalen we in de tekst dan ook als "45 procent van de leerlingen heeft deze vraag goed beantwoord".

<sup>18</sup> Als we de referentiegroepen onderling vergelijken, dan zijn er veel opgaven waarop p'-waarden significant tussen groepen verschillen. Voor het samenstellen van een 'top drie' van onderscheidende kenmerken is gekeken naar de effectgrootte (Cohen's d) per significant verschil tussen groepen.

De overige twee opgaven met grote verschillen tussen referentiegroepen <1F en 1F gaan beide over eenzelfde fragment over de Nederlandsche Bank. Bij een opgave over dit fragment is het de bedoeling een relatie te leggen tussen tekst en beeld (*begrijpen*, 2F). Bij de andere opgave gaat het erom het doel van de makers van een programma te verwoorden (*interpreteren*, 2F). Omdat de opgaven inhoudelijk gezien niet afwijken van overige opgaven in de luistertoets, lijkt vooral de inhoud van dit specifieke fragment bepalend te zijn voor het verschil tussen referentiegroepen.

Over het algemeen zijn verschillen tussen de referentiegroepen 1F en 2F kleiner dan tussen <1F en 1F. Van de drie opgaven (over drie verschillende fragmenten) die het meest onderscheidend zijn voor de groepen 1F en 2F, doen er twee een beroep op *begrijpen*. In de ene opgave moeten leerlingen hoofdzaken uit de tekst halen (*begrijpen* op niveau 1F). In de andere opgave moeten ze de betekenis van onbekende woorden (in dit geval: 'negatieve interest') afleiden uit de vorm, woordsoort, samenstelling of context van het fragment (niveau 2F). De derde onderscheidende opgave doet een beroep op *interpreteren* op niveau 1F. In deze opgave moeten leerlingen het doel van de makers van het programma verwoorden.

Samengevat is er geen specifiek tekst- of taakkenmerk dat het verschil tussen de groepen 1F en 2F verklaart. Onderscheidende opgaven doen zowel een beroep op (verschillende aspecten van) *begrijpen* als *interpreteren* op niveau 1F en 2F. Deze zijn gebaseerd op zowel audio- als videofragmenten. Wel behoren alle onderscheidende opgaven bij één van de twee type taken, namelijk *luisteren naar radio en televisie en gesproken tekst op internet* en niet op *luisteren naar instructie*.

### LUISTEREN IN HET KORT

Ruim 95 procent van de leerlingen beheerst tenminste het fundamentele niveau voor luisteren. 40 procent van de leerlingen beheerst ook het streefniveau. Daarmee behalen de leerlingen aan het einde van het basisonderwijs de ambitie die door de commissie Meijerink voor het fundamentele niveau gesteld werd (ten minste 85 procent van de leerlingen behaalt 1F). Dat geldt niet voor de ambitie voor het streefniveau (minimaal 65 procent 2F). De leerlingprestaties zijn gelijk voor de twee kenmerken van de taakuitvoering, *begrijpen* en *interpreteren*. Ook is er geen specifiek tekst- of taakkenmerk dat verschillen tussen de referentiegroepen verklaart.

### 3.2.3

#### Spreken

De spreekvaardigheid wordt in het referentiekader bepaald aan de hand van één taak: het houden van een monoloog. Daarbij worden er vijf kenmerken onderscheiden die bij het houden van een monoloog het referentieniveau bepalen. Dit zijn (zie ook tabel 3.2.3a):

- Samenhang
- Afstemming op doel
- Afstemming op publiek
- Woordgebruik en woordenschat
- Vloeiendheid, verstaanbaarheid en grammaticale beheersing

De uitvoering van de opdracht is in dit peilingsonderzoek als een randvoorwaardelijk kenmerk toegevoegd aan de beoordeling. Het maakt geen deel uit van het referentiekader.

Tabel 3.2.3a – Kenmerken van de taakuitvoering spreken (Staatsblad van het Koninkrijk der Nederlanden, 2010) samen met beoordelingsaspecten uit het beoordelingsmodel van het huidige peilingsonderzoek (Van Langen et al., 2017)

Kenmerk taakuitvoering	Aspect	Niveau 1F	Niveau 2F
<b>Uitvoering (Randvoorwaarde aan de spreekopdracht)</b>	<ul style="list-style-type: none"> <li>• Volledigheid</li> <li>• Inhoudelijke kwaliteit</li> </ul>		
<b>Samenhang</b>	<ul style="list-style-type: none"> <li>• Samenhang op macro- en mesoniveau: opbouw</li> <li>• Samenhang op microniveau: signaalwoorden</li> </ul>	<ul style="list-style-type: none"> <li>• Maakt eigen gedachtegang voor de luisteraar begrijpelijk, hoewel de structuur van de tekst nog niet altijd klopt.</li> </ul>	<ul style="list-style-type: none"> <li>• Kan een duidelijk verhaal houden met een samenhangende opsomming van punten en kan daarbij het belangrijkste punt duidelijk maken.</li> <li>• Gebruikt korte eenvoudige zinnen en verbindt deze door de juiste, eenvoudige voegwoorden en verbindingswoorden.</li> </ul>
<b>Afstemming op doel</b>	<ul style="list-style-type: none"> <li>• Spreekdoel tot uiting brengen</li> </ul>	<ul style="list-style-type: none"> <li>• Blijft trouw aan spreekdoel, soms met hulp van een ander.</li> </ul>	<ul style="list-style-type: none"> <li>• Geeft spreekdoel duidelijk vorm (instruerend, informatief, onderhoudend enz.) zodat het voor de luisteraar herkenbaar is.</li> </ul>
<b>Afstemming op publiek</b>	<ul style="list-style-type: none"> <li>• Taalgebruik afstemmen</li> <li>• Publiek boeien</li> <li>• Non-verbale communicatie</li> </ul>	<ul style="list-style-type: none"> <li>• Past het taalgebruik aan de luisteraar(s) aan.</li> <li>• Kan gebruik maken van ondersteunende materialen om een voorbereide presentatie beter aan het publiek over te brengen.</li> </ul>	<ul style="list-style-type: none"> <li>• Kan het verschil tussen formele en informele situaties hanteren.</li> <li>• Maakt de juiste keuze voor het register en het al dan niet hanteren van taalvariatie (dialect, jongerentaal).</li> <li>• Kan de luisteraar(s) boeien door middel van concrete voorbeelden en ervaringen.</li> </ul>
<b>Woordgebruik en woordenschat</b>	<ul style="list-style-type: none"> <li>• Woordgebruik</li> <li>• Woordenschat</li> </ul>	<ul style="list-style-type: none"> <li>• Beschikt over voldoende woorden om te praten over vertrouwde situaties en onderwerpen, maar zoekt nog regelmatig naar woorden en varieert niet veel in woordgebruik.</li> </ul>	<ul style="list-style-type: none"> <li>• Beschikt over voldoende woorden om zich te kunnen uiten.</li> <li>• Het kan soms nodig zijn een omschrijving te geven van een onbekend woord.</li> </ul>
<b>Vloeiendheid, verstaanbaarheid en grammaticale beheersing</b>	<ul style="list-style-type: none"> <li>• Vloeiendheid</li> <li>• Uitspraak en intonatie</li> <li>• Grammatica: vervoeging</li> <li>• Grammatica: verbuiging</li> <li>• Grammatica: zinsbouw</li> </ul>	<ul style="list-style-type: none"> <li>• De uitspraak is duidelijk genoeg om de spreker te kunnen volgen, ondanks een eventueel accent, verkeerde intonatie, onduidelijke articulatie en/of haperingen.</li> <li>• Redelijk accuraat gebruik van eenvoudige zinsconstructies.</li> <li>• Houding, intonatie en mimiek ondersteunen het gesprokene.</li> <li>• Pauzes, valse starts en herformuleringen komen af en toe voor.</li> </ul>	<ul style="list-style-type: none"> <li>• De uitspraak is duidelijk verstaanbaar, ondanks een eventueel accent, af en toe een verkeerd uitgesproken woord en/of haperingen.</li> <li>• Vertoont een redelijke grammaticale beheersing.</li> <li>• Aarzelingen en fouten in zinsbouw zijn eigen aan gesproken taal en komen dus voor, maar worden zonnodig hersteld.</li> <li>• Is goed te volgen en kan zich gemakkelijk uitdrukken.</li> <li>• Vloeiendheid kan minder zijn als er nagedacht moet worden over de grammaticale vorm, de te kiezen woorden en herstel van fouten.</li> </ul>

In het peilingsonderzoek hielden de leerlingen een monoloog in de vorm van een vlog over een fictieve gebeurtenis, namelijk het winnen van een straatprijs. Er is gekozen voor een vlog, omdat dit een herkenbare spreek situatie is voor de leerling. De spreekvaardigheid in dit onderzoek is met één taak gemeten. Dat heeft gevolgen voor de generaliseerbaarheid van de resultaten. Deze moeten geïnterpreteerd worden in de context van de gekozen spreek situatie (de vlog). De resultaten kunnen niet worden gegeneraliseerd naar andere spreek situaties (zoals een verhaal vertellen aan de hand van een aantal plaatjes) en daarmee naar een algemeen beeld van de spreekvaardigheid einde basisonderwijs.

Om de taak uit te kunnen voeren, kregen leerlingen van de toetsleider voorafgaand aan de taak een gestandaardiseerde instructie, samen met een compilatievideo met voorbeelden van vlogs. Zo werd de voorkennis van de leerlingen geactiveerd en leerden leerlingen die nog niet bekend waren met vlogs, wat een vlog inhoudt. Na deze instructie kreeg elke leerling maximaal vijftien minuten de tijd om een vlog voor te bereiden met een zogenaamde 'spreekhulp', zie figuur 3.2.3a.

Figuur 3.2.3a – Spreekhulp bij het maken van een vlog

SPREEKHULP		Een vlog over de loterij winnen	
<p>Wat voor mensen kijken er naar mijn vlog? Wie zijn mijn volgers?</p> <p>Kinderen van mijn leeftijd</p>	<p>JE BESPREEKT IN ELK GEVAL:</p> <p>Wat is er gebeurd bij jou in de buurt?</p>	<p>Je hebt 15 minuten de tijd om deze spreekhulp in te vullen. Wil je iets vragen? Steek dan je vinger op.</p>	<p>STEEKWOORDEN VOOR DE VLOG</p>
<p>Wat wil ik met mijn vlog bereiken?</p> <p>Mijn volgers vertellen wat ik meemaak en wat ik daarvan vind</p>	<p>Wat voor mensen wonen er in die straat?</p>		<p>Gebeurtenis in de buurt</p>
<p>Wat is de belangrijkste boodschap van mijn vlog?</p> <p>_____</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>Hoe reageerden deze mensen?</p>		<p>Soort mensen in de buurt</p>
<p>Hoe ga je de vlog afsluiten?</p> <p>_____</p> <p>_____</p> <p>_____</p> <p>_____</p>	<p>Wat vind jij van deze gebeurtenis?</p>		<p>Reacties van de mensen</p>
	<p>Wat gaat er veranderen in jouw buurt, denk je?</p>		<p>Eigen mening over gebeurtenis</p>
	<p>Stel dat jij zo veel geld zou winnen. Wat zou je dan doen voor jezelf?</p>		<p>Verandering in de buurt?</p>
	<p>En wat zou je doen voor een ander?</p>		<p>Wat zou ik doen, voor mezelf en voor anderen? Waarom?</p>
	<p>Waarom zou je deze dingen doen?</p>		
	<p>Wat zijn voordelen van het winnen van veel geld?</p>		<p>Voor- en nadelen van veel geld</p>
	<p>En wat zijn nadelen?</p>		
	<p>Vind je het belangrijk om veel geld te hebben?</p>		<p>Geld: belangrijk? Waarom?</p>
	<p>Waarom vind je dat?</p>		

De spreekhulp zorgde ervoor dat leerlingen dezelfde uitgangspunten gebruikten bij het maken van de vlog. Ook mochten leerlingen op de spreekhulp in steekwoorden noteren wat ze bij elk van de te bespreken onderwerpen wilden zeggen. Tijdens het opnemen van de vlog mochten ze het rechtergedeelte van de spreekhulp (zie figuur 3.2.3a) gebruiken als spiekbriefje. De vlogs duurden drie tot vijf minuten.

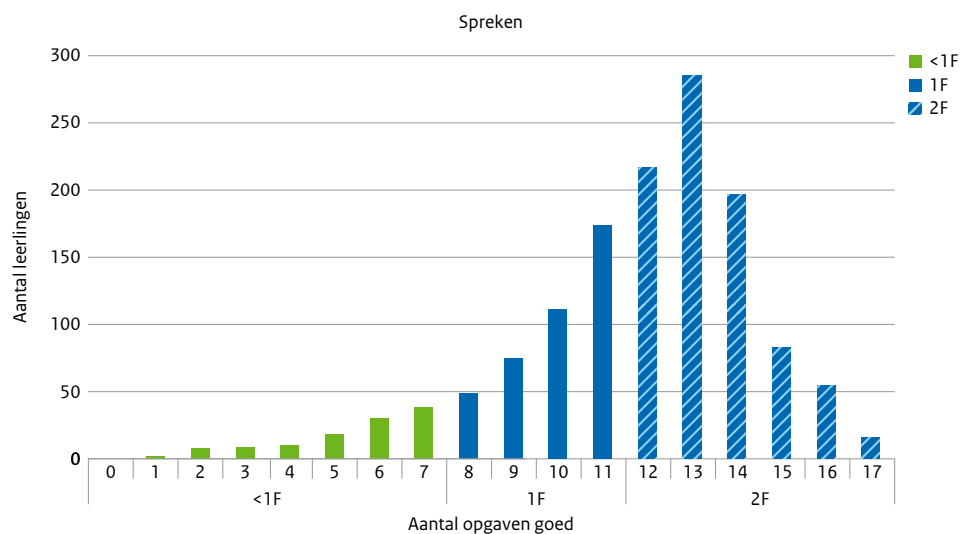


De leerlingen werden achteraf aan de hand van de gemaakte video's beoordeeld met een gestandaardiseerd beoordelingsmodel (Van Langen et al, 2017). De leerlingen zijn daarbij op de vijf kenmerken uit het referentiekader beoordeeld en op een toegevoegd kenmerk: de uitvoering van de monoloog (zie tabel 3.2.3a).

Bij het berekenen van de totaalscore op basis van deze beoordelingen wogen de meer technische kenmerken, zoals *woordgebruik en woordenschat* en *vloeiendheid, verstaanbaarheid en grammaticale beheersing*, minder zwaar mee (samen voor 25 procent). Deze kenmerken worden meer beschouwd als middelen die bijdragen aan de spreekvaardigheid, maar daar geen centrale, inhoudelijke rol in spelen. Leerlingen konden een maximale score van 17 punten behalen.

Figuur 3.2.3b laat zien hoeveel leerlingen een score 0 tot en met 17 hebben behaald. Geen enkele leerling scoort 0. 16 leerlingen hebben de maximale score behaald. 92 procent van de leerlingen beheerst minimaal het fundamentele referentieniveau voor spreken. Dat is ruim boven het ambitieniveau van de commissie Meijerink. Ook beheerst ruim 62 procent het streefniveau. Daarmee is de ambitie ten aanzien van het streefniveau, minimaal 65 procent op niveau 2F, ook bijna behaald.

Figuur 3.2.3b – Scores op de toets spreken (n=1365)

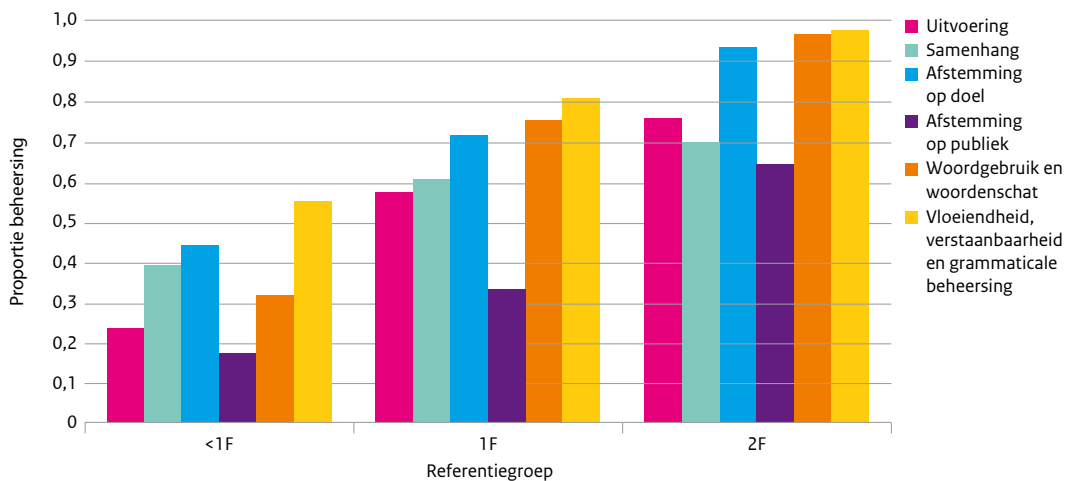


### De afzonderlijke kenmerken voor spreken

Voor de totale groep leerlingen geldt dat zij gemiddeld de hoogste beoordeling krijgen voor het kenmerk *vloeiendheid, verstaanbaarheid en grammaticale beheersing*. Gemiddeld kregen ze de laagste beoordeling voor *afstemming op publiek*. Blijkbaar is dit laatste kenmerk een lastig onderdeel bij het houden van een monoloog.

Figuur 3.2.3c laat per referentiegroep zien waar leerlingen gemiddeld hoog en waar ze gemiddeld laag op beoordeeld worden. We zien dat het relatieve beheersingsniveau van de verschillende taakkenmerken hetzelfde is voor alle referentiegroepen, op één kenmerk na: *woordgebruik en woordenschat* van leerlingen die het niveau 1F nog niet beheersen. Deze groep leerlingen heeft hier, in vergelijking met de andere kenmerken, beduidend (en significant) meer moeite mee dan de overige twee groepen. Daarnaast maakt de figuur duidelijk dat er grotere verschillen bestaan tussen leerlingen die het fundamentele niveau nog niet beheersen, dan tussen leerlingen die 1F of hoger scoren. In de groep leerlingen met niveau 2F zijn de verschillen tussen de afzonderlijke kenmerken onderling het kleinst.

Figuur 3.2.3c – Gemiddelde proportie beheersing per kenmerk van de taakuitvoering (n=1365)



### Wat onderscheidt leerlingen presterend op de verschillende referentieniveaus van elkaar?

Ieder kenmerk van de taakuitvoering is beoordeeld aan de hand van meerdere aspecten. Zo komt de score voor het kenmerk *samenhang* bijvoorbeeld tot stand op basis van scores op de beoordelingsaspecten *samenhang op macro- en mesoniveau* en *samenhang op microniveau*. Tabel 3.2.3a (p. 75) geeft een overzicht van de beoordelingsaspecten per taakkenmerk, zoals deze zijn opgenomen in het beoordelingsmodel van spreken. Als we hier op inzoomen, zien we welke specifieke beoordelingsaspecten leerlingen die het fundamentele niveau 1F niet beheersen met name onderscheiden van leerlingen die dit wel doen. Datzelfde geldt voor het onderscheiden van aspecten die vooral bijdragen aan het verschil tussen leerlingen op niveau 1F of 2F.

Het verschil tussen de leerlingen die het fundamentele niveau voor spreekvaardigheid wel of nog niet beheersen (1F versus <1F) is vooral zichtbaar bij de prestaties op de inhoudelijke kwaliteit van de uitvoering (zie ook kader op de volgende pagina). Daarnaast zijn leerlingen die het fundamentele niveau beheersen ook vollediger in hun uitvoering en krijgen ze hogere scores voor woordgebruik.

Leerlingen die 1F en leerlingen die 2F beheersen, verschillen met name van elkaar in de mate waarin ze het publiek weten te boeien: zowel het aspect *publiek boeien* als het aspect *non-verbale communicatie* worden bij leerlingen met niveau 2F beter beoordeeld. Ook de mate waarin de leerlingen het spreekdoel tot uiting weten te brengen verschilt tussen leerlingen van referentiegroep 1F en 2F. Hier geldt eveneens dat leerlingen in de groep 2F significant hogere scores behalen voor deze kenmerken dan leerlingen in de groep 1F.

Op de volgende pagina's staat een aantal citaten die de verschillen tussen leerlingen in de referentiegroepen <1F, 1F en 2F illustreren. Deze voorbeelden zijn geselecteerd uit opnames van leerlingen die op de onderscheidende aspecten een score behaalden die typerend is voor hun referentiegroep (voor meer citaten, zie Van Langen et al, 2017).

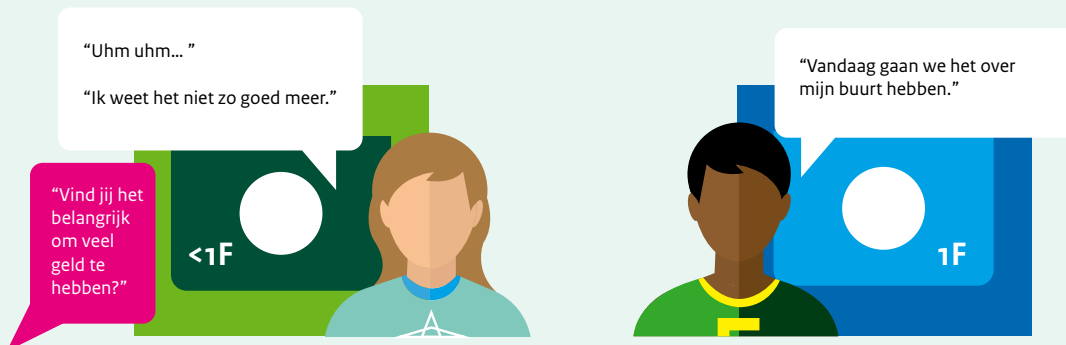
### Uitvoering: **kwaliteit van de inhoud**

De uitwerking van leerlingen die niveau 1F nog niet behalen kenmerkt zich door de weinig complexe of originele inhoud. Daarnaast zit er weinig chronologie in de monoloog; de leerling noemt puntsgewijs datgene op dat in de spreekhulp staat opgeschreven. De uitvoering van leerlingen met niveau 1F kenmerkt zich door een wat meer complexe inhoud, bijvoorbeeld door nuances aan te brengen in een eigen mening of een tegenstelling te markeren. Ook gebruikt de 1F-leerling treffendere en originelere voorbeelden dan de <1F-leerling.



### Uitvoering: **volledigheid**

Leerlingen die het niveau 1F nog niet beheersen, lukt het meestal niet de opdracht volledig uit te voeren zonder hulp van de toetsleider. De toetsleider interveenieert met vragen aan de leerling, waardoor deze verder vertelt. In de vlog laat de <1F-leerling het merendeel van de beoogde taalhandelingen (beschrijven van een situatie, proces en mening) zien; de 1F-leerling daarentegen voert alle taalhandelingen uit.



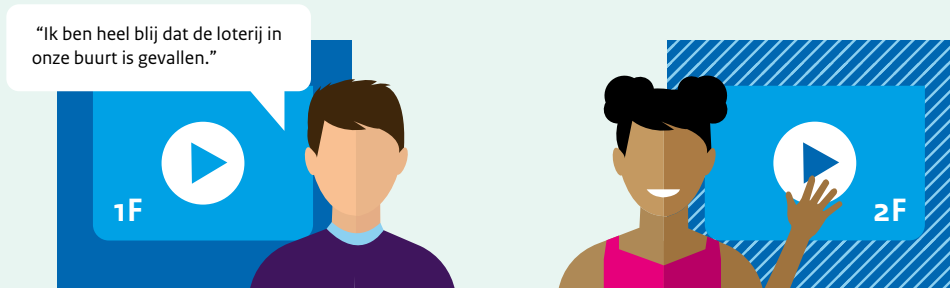
### Woordenschat en woordgebruik: **woordgebruik**

De leerlingen die niveau 1F niet behalen tonen weinig variatie in woordgebruik. Voorbeelden hiervan zijn het herhalen van bepaalde frases. We zien dat leerlingen op niveau 1F in dat woordgebruik wat meer variatie en nuances aanbrengen bijvoorbeeld op ‘leuk’, ‘blij’ en ‘niet leuk’.



### Afstemming op het publiek: **non-verbale communicatie**

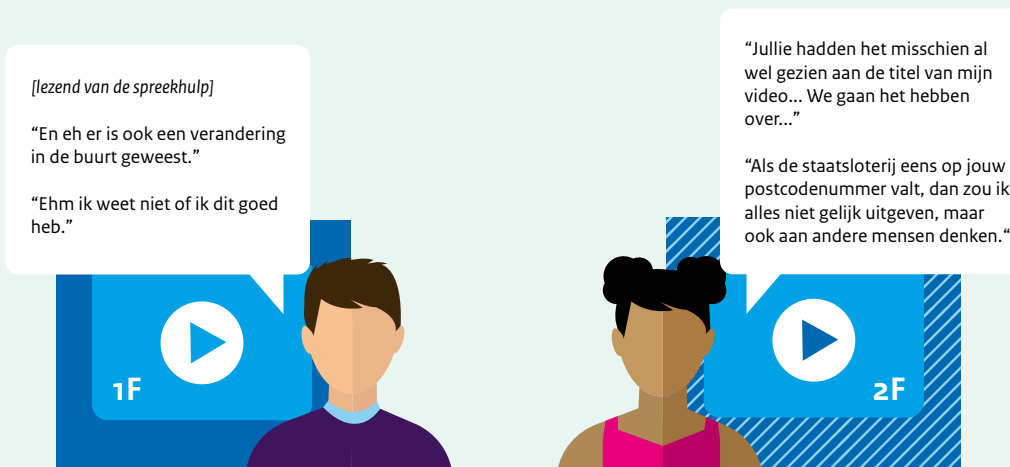
Een van de verschillen tussen leerlingen met niveau 1F en 2F als het gaat om de taakuitvoering is de mate waarin ze non-verbale communicatie toepassen: leerlingen op niveau 2F kijken in de camera en maken contact met het publiek, bijvoorbeeld door te zwaaien. De 'typische' 1F-leerling kijkt vaker weg van de camera. Daarnaast past het non-verbale gedrag (bijvoorbeeld een neutrale blik) van de 1F-leerling niet altijd bij de verbale boodschap.



"Ik ben heel blij dat de loterij in onze buurt is gevallen."

### Afstemming op publiek: **het publiek boeien**

Ook in de manier waarop leerlingen het publiek weten te boeien, zijn er verschillen tussen leerlingen op niveau 1F en op niveau 2F. Bij leerlingen op niveau 2F zien we vaker expliciete uitingen om de kijker mee te nemen in hun vlog. Ze doen dat bijvoorbeeld door hun eigen rol als vlogger te expliciteren, maar ook door ervaringen te delen op een levensechte manier. 1F-leerlingen vallen vaker uit hun rol en vragen bevestiging van de toetsleider. Ook spreken leerlingen op niveau 2F kijkers vaker direct aan, soms zelfs op hun gevoel.



[lezend van de spreekhulp]

"En eh er is ook een verandering in de buurt geweest."

"Ehm ik weet niet of ik dit goed heb."

"Jullie hadden het misschien al wel gezien aan de titel van mijn video... We gaan het hebben over..."

"Als de staatsloterij eens op jouw postcodenummer valt, dan zou ik alles niet gelijk uitgeven, maar ook aan andere mensen denken."

### Afstemming op doel

Leerlingen op niveau 2F brengen het spreekdoel veel explicieter tot uiting in hun vlog dan leerlingen op niveau 1F. 2F-leerlingen geven vaker expliciet aan wat hun spreekdoel is en hebben vaker een duidelijke boodschap voor de kijkers van hun vlog. Ook sluiten leerlingen op niveau 2F vaker hun vlog zelf af, met bijbehorende boodschap. De leerlingen op 1F-niveau hebben hulp nodig van de toetsleider bij het afsluiten van de vlog.



"...uhm, het moest drie minuten duren toch?"

"Nou, als je klaar bent mag je afronden hoor."

"Oke, dit was het einde van de video. Doe!"

"Bedankt voor het kijken en doe een like als je ook vindt dat vrienden en familie belangrijker zijn dan het winnen van veel geld."

"...maar ik wilde jullie iets vertellen over wat er bij mij in de straat is gebeurd. (...) en daarom maak ik deze video."

### SPREKEN IN HET KORT

92 procent van de leerlingen beheerst ten minste het fundamentele referentieniveau voor spreken. 62 procent van de leerlingen beheerst ook het streefniveau. Net als bij het subdomein luisteren, wordt daarmee de ambitie voor het fundamenteel niveau ruimschoots behaald. Ook de ambitie ten aanzien van het streefniveau wordt op het subdomein spreken, in tegenstelling tot de andere subdomeinen, bijna behaald. Voor de totale groep leerlingen geldt dat, van alle taakkenmerken, de *afstemming op het publiek* het laagst wordt beoordeeld. Leerlingen die het fundamenteel niveau nog niet beheersen, hebben gemiddeld meer moeite met *woordgebruik en woordenschat* dan leerlingen die al wel op fundamenteel niveau presteren.

Als we verder inzoomen op de prestaties en kijken naar de beoordeelde aspecten (onderliggend aan de kenmerken van de taakuitvoering), dan zien we dat vlogs van leerlingen die het fundamenteel niveau nog niet beheersen zich onderscheiden van vlogs van 1F-leerlingen door:

- weinig complexe inhoud;
- onvolledigheden en hulpbehoefendheid in de uitvoering;
- weinig variatie in woordgebruik.

Vergelijken we de vlogs van 2F-leerlingen met die van 1F-leerlingen, dan valt op dat 2F-leerlingen:

- meer (passende) nonverbale communicatie toepassen;
- het publiek expliciet betrekken in hun vlog;
- vaker een duidelijke boodschap hebben voor het publiek.

### 3.2.4

#### Gesprekken

In het referentiekader wordt het subdomein gesprekken uiteengelegd in twee taken: *Deelnemen aan discussie en overleg* en *informatie uitwisselen*. In dit onderzoek is de gespreksvaardigheid van leerlingen beoordeeld aan de hand van één taak, waarin *deelnemen aan discussie en overleg* centraal stond<sup>19</sup>. Net als bij spreken geldt ook hier dat de resultaten van dit onderzoek dan ook geïnterpreteerd moeten worden binnen de context van de gekozen gesprekssituatie.

De leerlingen uit groep 8 voerden in drietallen gesprekken over het organiseren van een actie om geld in te zamelen voor een goed doel. Daarnaast kregen ze een (fictief) bedrag van vijftig euro. Daarmee konden ze de voorbereiding van die actie financieren. De leerlingen moesten gezamenlijk afspreken hoe ze dit geld gingen besteden en wat de bijbehorende planning van het project was. Om hen daarbij te helpen, kregen ze als groepje een kaart. Daarop stonden de te bespreken onderwerpen. (zie figuur 3.2.4a).

Figuur 3.2.4a – Gesprekskaart met onderwerpen

### GESPRESKSONDERWERPEN

1	Voor welk goed doel gaan jullie geld inzamelen en waarom?
2	Welke acties gaan jullie organiseren en hoe gaan jullie de € 50,- besteden?
3	Wanneer gaan jullie de acties uitvoeren?

Daarnaast kreeg iedere leerling een eigen gesprekskaart. Daarop stond informatie over de aard van het goede doel, het soort inzamelactie en de planning die specifiek voor de leerling bedoeld was. Er waren in het gesprek drie leerlingrollen, die elk voor bepaalde momenten in het gesprek een eigen inbreng vereisten. Tabel 3.2.4a geeft een overzicht van de verschillende leerlingrollen.

<sup>19</sup> Deze keuze, om minimaal de taak *discussie en overleg* te beoordelen, is gemaakt op basis van de domeinbeschrijving (SLO, 2016). Deze is mede ingegeven door de wens om een zinvolle en authentieke taak af te nemen.

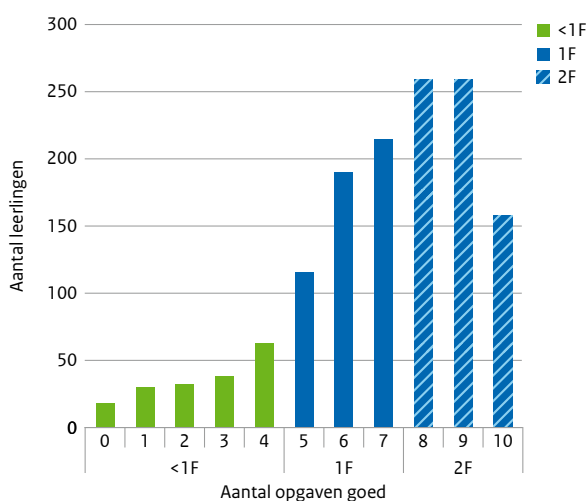
Tabel 3.2.4a – Informatie per leerlingrol gesprekken

	Keuze goed doel	Soort inzamelactie	Planning
Leerling 1	dieren	[geen informatie]	zo snel mogelijk
Leerling 2	[geen informatie]	iets maken en verkopen	de tijd nemen
Leerling 3	zieke kinderen	iets organiseren	[geen informatie]

Door de informatie op de gesprekskaarten werd de leerling gestimuleerd om op minimaal twee momenten in het gesprek actief iets in te brengen. Daarnaast mocht de leerling ook meepraten over andere onderwerpen die tijdens het gesprek aan bod kwamen. Voor het goede verloop van het gesprek kon de toetsleider als het gesprek vastliep, ingrijpen volgens een vooraf ingevuld gespreksschema met stuurvragen. Alle gesprekken werden opgenomen op video. Deze zijn achteraf, op individueel niveau, beoordeeld met een gestandaardiseerd beoordelingsmodel (zie Van Langen et al., 2017). Een leerling kon voor het gesprek dat hij had gevoerd een maximale score van 10 behalen en een minimale score van 0 punten.

In figuur 3.2.4b staat de verdeling van de leerlingen over de scores voor het subdomein gesprekken weergegeven, evenals de indeling naar referentieniveaus. Zo'n 87 procent van de leerlingen beheerst minimaal het fundamentele niveau 1F. Dit is boven het ambitieniveau van 85 procent van de commissie Meijerink. Het ambitieniveau van 65 procent ten aanzien van het streefniveau wordt niet behaald: 49 procent van de leerlingen beheerst het streefniveau 2F voor gespreksvaardigheid.

Figuur 3.2.4b – Scoreverdeling op de toets gesprekken (n=1380)



### De afzonderlijke kenmerken voor gesprekken

De gesprekstaak is voor iedere individuele leerling beoordeeld op een aantal kenmerken (zie ook tabel 3.2.4b):

- Beurten nemen en bijdragen aan samenhang
- Afstemming op doel
- Afstemming op gesprekspartner
- Woordgebruik en woordenschat en vloeiendheid, verstaanbaarheid en grammaticale beheersing<sup>20</sup>

<sup>20</sup> In het referentiekader (Staatsblad van het Koninkrijk der Nederlanden, 2010) zijn woordgebruik en woordenschat versus vloeiendheid, verstaanbaarheid en grammaticale beheersing twee aparte kenmerken van de taakuitvoering. In het beoordelingsmodel van de huidige peiling zijn deze kenmerken samengenomen.

De uitvoering van de opdracht is, net als bij het subdomein spreken, als een randvoorwaardelijk kenmerk toegevoegd aan de beoordeling (gemeten als: Kwaliteit van de inhoud).

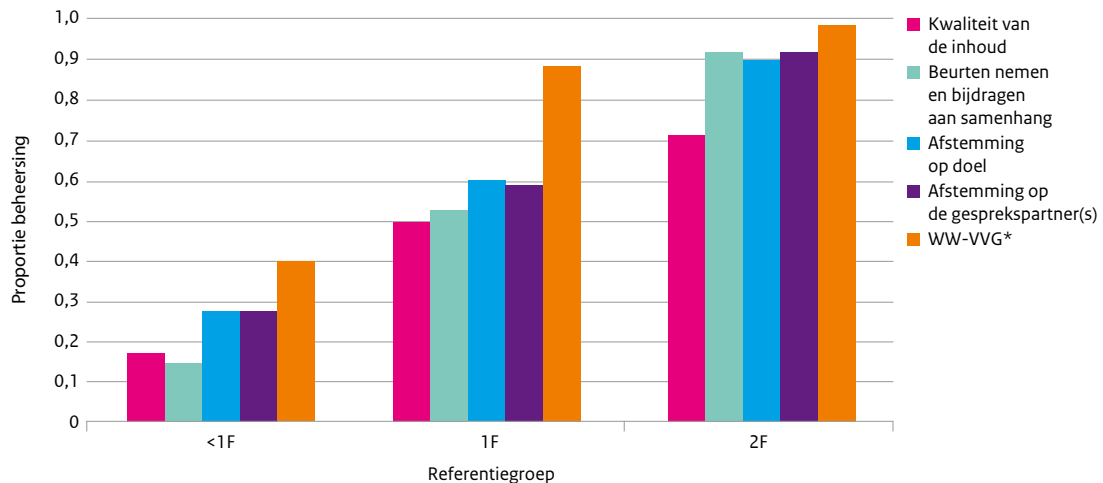
Tabel 3.2.4b – Kenmerken van de taakuitvoering gesprekken (Staatsblad van het Koninkrijk der Nederlanden, 2010) samen met beoordelingsaspecten uit het beoordelingsmodel van het huidige peilingsonderzoek (Van Langen et al., 2017)

Kenmerk taakuitvoering	Aspect	Niveau 1F	Niveau 2F
<b>Uitvoering (Randvoorwaardelijk voor gesprekstaak)</b>	<ul style="list-style-type: none"> <li>• Kwaliteit van de inhoud</li> </ul>		
<b>Beurten nemen en bijdragen aan samenhang</b>	<ul style="list-style-type: none"> <li>• Gesprek onderhouden</li> <li>• Initiatief in beurten nemen</li> <li>• Passendheid van beurten nemen</li> <li>• Inbreng gesprekspartners integreren</li> </ul>	<ul style="list-style-type: none"> <li>• Kan een kort gesprek beginnen, gaande houden en beëindigen.</li> </ul>	<ul style="list-style-type: none"> <li>• Kan de juiste frase gebruiken om aan het woord te komen.</li> <li>• Kan een reactie uitstellen totdat hij de bijdrage van de ander geïnterpreteerd en beoordeeld heeft.</li> </ul>
<b>Afstemming op doel</b>	<ul style="list-style-type: none"> <li>• Doelgerichtheid</li> </ul>	<ul style="list-style-type: none"> <li>• Kan gesprekken voeren om informatie en meningen uit te wisselen, uitleg of instructie te geven en te volgen.</li> <li>• Herkent gespreksituaties en kan passende routines gebruiken.</li> </ul>	<ul style="list-style-type: none"> <li>• Kan het eigen gespreksdoel tot uitdrukking brengen.</li> <li>• Kan doelgericht doorvragen om de gewenste informatie te verwerven.</li> </ul>
<b>Afstemming op de gesprekspartner(s)</b>	<ul style="list-style-type: none"> <li>• Taalgebruik afstemmen</li> <li>• Non-verbale communicatie</li> <li>• Actief luistergedrag</li> </ul>	<ul style="list-style-type: none"> <li>• Kan de gesprekspartner(s) redelijk volgen tenzij ze voor onverwachte wendingen in het gesprek zorgen.</li> <li>• Kan woorden ondersteunen met non-verbaal gedrag.</li> </ul>	<ul style="list-style-type: none"> <li>• Kan het spreekdoel van anderen herkennen en reacties schatten.</li> <li>• Kan het verschil tussen formele en informele situaties hanteren.</li> <li>• Maakt de juiste keuze voor het register en het al dan niet hanteren van taalvariatie (dialect, jongerentaal).</li> </ul>
<b>Woordgebruik en woordenschat</b>	<ul style="list-style-type: none"> <li>• Woordgebruik</li> <li>• Woordenschat</li> </ul>	<ul style="list-style-type: none"> <li>• Beschikt over voldoende woorden om te praten over vertrouwde situaties en onderwerpen, maar zoekt nog regelmatig naar woorden en varieert niet veel in woordgebruik.</li> </ul>	<ul style="list-style-type: none"> <li>• Beschikt over voldoende woorden om zich te kunnen uiten.</li> <li>• Het kan soms nodig zijn een omschrijving te geven van een onbekend woord.</li> </ul>
<b>Vloeiendheid, verstaanbaarheid en grammaticale beheersing</b>	<ul style="list-style-type: none"> <li>• Vloeiendheid</li> <li>• Uitspraak en intonatie</li> <li>• Grammatica: vervoeging</li> <li>• Grammatica: verbuiging</li> <li>• Grammatica: zinsbouw</li> </ul>	<ul style="list-style-type: none"> <li>• De uitspraak is duidelijk genoeg om de spreker te kunnen volgen, ondanks een eventueel accent, verkeerde intonatie, onduidelijke articulatie en/of haperingen.</li> <li>• Redelijk accuraat gebruik van eenvoudige zinsconstructies.</li> </ul>	<ul style="list-style-type: none"> <li>• De uitspraak is duidelijk verstaanbaar, ondanks een eventueel accent, af en toe een verkeerd uitgesproken woord en/of haperingen.</li> <li>• Vertoont een redelijke grammaticale beheersing.</li> <li>• Aarzelingen en fouten in zinsbouw zijn eigen aan gesproken taal en komen dus voor, maar worden zo nodig hersteld.</li> </ul>

De scores op de kenmerken *woordgebruik en woordenschat* en *vloeiendheid, verstaanbaarheid en grammaticale beheersing* zijn bij elkaar opgeteld. Deze telden voor 25 procent mee in de totaalscore. De kenmerken *uitvoering: kwaliteit van de inhoud, beurten nemen* en *afstemming op doel* wogen zwaarder mee dan de overige kenmerken. Ook werd actief luistergedrag als aspect van het kenmerk *afstemming op de gesprekspartner* zwaarder meegewogen.

We hebben berekend hoe leerlingen gemiddeld presteren op elk kenmerk van de taakuitvoering. Dat hebben we gedaan door voor elk kenmerk de proportie beheersing uit te rekenen; ofwel de gemiddelde hoogte van het oordeel op het betreffende kenmerk ten opzicht van de maximale score (zie figuur 3.2.4c).

Figuur 3.2.4c – Gemiddelde proportie beheersing per kenmerk van de taakuitvoering (n=1380)



NB. WW-VVG = Woordgebruik en woordenschat en Vloeiendheid, verstaanbaarheid en grammaticale beheersing.

Voor de totale groep leerlingen geldt dat zij gemiddeld de hoogste beoordeling krijgen voor *woordgebruik en woordenschat en vloeiendheid, verstaanbaarheid en grammaticale beheersing*. Gemiddeld krijgen zij de laagste beoordeling op *kwaliteit van de inhoud*. Deze volgorde in beheersingsniveau is voor de afzonderlijke referentiegroepen min of meer hetzelfde. Alleen voor de groep leerlingen die het niveau 2F beheersen, geldt dat het kenmerk *kwaliteit van de inhoud* significant lager wordt beoordeeld dan alle andere kenmerken. Bij de andere twee referentiegroepen is er geen verschil in de beoordeling van de kenmerken *kwaliteit van de inhoud* en *beurten nemen en bijdragen aan samenhang*, maar worden beide kenmerken wel significant lager beoordeeld dan de overige drie kenmerken van de taakuitvoering.

### Wat onderscheidt leerlingen presterend op de verschillende referentieniveaus van elkaar?

Wat maakt nu vooral het verschil tussen de groepen <1F en 1F en de groepen 1F en 2F? Dit kunnen we achterhalen door naar de scores op de onderliggende beoordelingsaspecten bij de kenmerken van de taakuitvoering te kijken. Bij het kenmerk *afstemming op de gesprekspartner* gaat het om bijvoorbeeld drie aspecten, namelijk *taalgebruik afstemmen*, *non-verbale communicatie* en *actief luistergedrag* (zie ook tabel 3.2.4b). De scores op de onderliggende beoordelingsaspecten vormen samen de score van het taakkenmerk.

Het verschil tussen de groepen <1F en 1F is vooral dat leerlingen die het fundamentele niveau 1F beheersen, beter het gesprek gaande kunnen houden dan leerlingen die nog niet voldoen aan dit niveau. Ook nemen zij meer initiatief in het beurten nemen dan leerlingen die het fundamentele niveau nog niet beheersen. Beide aspecten zijn onderdeel van het taakkenmerk *beurten nemen en bijdragen aan samenhang*. Verder weten zij beter af te stemmen op het doel; leerlingen met niveau 1F krijgen een hogere beoordeling op het aspect *doelgerichtheid* dan leerlingen die het niveau 1F nog niet beheersen<sup>21</sup>. *Doelgerichtheid* is zowel een vaardigheid die de groep <1F onderscheidt van 1F, als die de groep 1F onderscheidt van 2F. Hoe hoger het referentieniveau, hoe hoger de gemiddelde score op dit aspect. Dit geldt ook voor het *onderhouden van het gesprek* (aspect bij het taakkenmerk *beurten nemen en bijdragen aan samenhang*): ook op dit aspect nemen gemiddelde scores toe, als het referentieniveau hoger wordt. Ten slotte vertonen leerlingen met niveau 2F meer *actief luistergedrag* dan leerlingen met niveau 1F (aspect bij het taakkenmerk *afstemming op de gesprekspartner(s)*).

<sup>21</sup> Ook verschilden de groepen <1F en 1F van elkaar in de kwaliteit van de inhoud van hun uitingen. Echter, omdat dit beoordelingsaspect relatief vaak als 'niet te beoordelen' werd beschouwd in de <1F-groep, is het niet opgenomen bij de citaten op de volgende bladzijde.



Hieronder staat een aantal citaten die de verschillen in gespreksvaardigheden tussen leerlingen in de referentiegroepen <1F (groen), 1F (blauw) en ≥2F (blauw gearceerd) illustreren. Deze citaten zijn geselecteerd uit opnames van leerlingen die binnen referentiegroepen typerende scores behalen op de onderscheidende aspecten.



**Gesprek onderhouden:** Leerlingen op niveau 2F nemen veel vaker initiatief om het gesprek zelfstandig te starten, sturen en af te sluiten. De bijdrage aan het gesprek van leerlingen op 1F kenmerkt zich vooral door het reageren op en/of kopiëren van de inbreng van anderen. Leerlingen die het 1F-niveau nog niet beheersen, houden zich afzijdig in het gesprek. Ze brengen alleen iets in als ze direct iets gevraagd wordt. Soms ondermijnt de inbreng van de leerling op <1F zelfs het gesprek, omdat deze op dat moment niet passend is.

**Doelgerichtheid:** Leerlingen op niveau 2F nemen initiatief om het gezamenlijke gespreksdoel te behalen door de gesprekspartners te betrekken bij het gesprek, bijvoorbeeld met behulp van vragen. Ook de leerling op niveau 1F doet actief mee om het gespreksdoel te behalen, maar wel vooral vanuit zijn/haar rol met behulp van de gesprekskaart. Leerlingen op niveau <1F leveren geen actieve bijdrage om het gespreksdoel te behalen.

**Initiatief in beurten nemen:** <1F-leerlingen en 1F-leerlingen onderscheiden zich naast bovengenoemde nog in de mate waarin er initiatief wordt genomen in het beurten nemen. Waar leerlingen met niveau 1F nog wel zelf initiatief nemen om iets toe te voegen aan het gesprek, gebeurt dit zelden bij leerlingen die het 1F-niveau nog niet beheersen.

**Actief luistergedrag:** Leerlingen met niveau 1F en niveau 2F verschillen ten slotte van elkaar in de mate waarin actief luistergedrag wordt getoond. Leerlingen met niveau 2F hebben – meer dan 1F-leerlingen - een actieve zithouding, maken oogcontact met anderen en reageren bevestigend (bijvoorbeeld door te knikken) tijdens het gesprek.

### GESPREKKEN IN HET KORT

87 procent van de leerlingen beheerst minimaal het fundamentele referentieniveau voor gesprekken. 49 procent van de leerlingen beheerst ook het streefniveau. Daarmee wordt de ambitie ten aanzien van het fundamentele niveau behaald. Dat geldt niet voor het streefniveau, al zijn er wel veel leerlingen die dit niveau *nét* niet halen. Gemiddeld krijgt de totale groep leerlingen de laagste beoordeling voor de *kwaliteit van de inhoud*. Alleen 2F-leerlingen scoren op dit kenmerk van de taakuitvoering significant lager dan op alle andere kenmerken. Leerlingen die niet op streefniveau presteren, hebben met de kenmerken *kwaliteit van de inhoud* en *beurten nemen en bijdragen aan samenhang* evenveel moeite, meer dan met de overige kenmerken van de taakuitvoering.

Als we verder inzoomen op de prestaties en kijken naar de beoordeelde aspecten onderliggend aan de taakkenmerken, dan zien we dat als het referentieniveau toeneemt, leerlingen

- a) vaker het initiatief nemen om het gesprek te starten, sturen of af te sluiten en
- b) vaker initiatief nemen om het gespreksdoel te behalen, door gesprekspartners te betrekken bij het gesprek.

Daarnaast valt op dat 1F-leerlingen soms zelf initiatief nemen om iets toe te voegen aan een gesprek, terwijl leerlingen die het 1F-niveau nog niet beheersen dit zelden doen. Ten slotte vertonen 2F-leerlingen, meer dan 1F-leerlingen, actief luistergedrag, bijvoorbeeld door oogcontact te maken met anderen.

## 3.3 Trends

Hoe heeft de mondelinge taalvaardigheid van leerlingen in groep 8 zich ontwikkeld? Presteren de huidige basisschoolleerlingen beter dan basisschoolleerlingen een aantal jaren terug? Om dit na te gaan, is de luister- en spreekvaardigheid van de leerlingen uit dit peilingsonderzoek vergeleken met de luister- en spreekvaardigheid van leerlingen uit vergelijkbare peilingsonderzoeken die in respectievelijk 2007 en 2010 zijn uitgevoerd (Krom, Van Berkel, Van der Schoot, Sijstra, Hemker, & Marsman 2011; Krämer, Kuhlemeier, Knoop, Hemker, & Van Weerden, 2014). Om deze vergelijking te maken, zijn een aantal luisteritems uit het peilingsonderzoek van 2007 ook meegenomen in het huidige onderzoek. De scores op deze zogenaamde ‘ankeritems’ kunnen met elkaar vergeleken worden. Voor een vergelijking van de spreekvaardigheid tussen de twee onderzoeken, maakten een aantal leerlingen uit het huidige peilingsonderzoek een spreekopdracht uit het peilingsonderzoek van 2010 (een ‘ankertoets’). Die is op basis van het bijbehorende beoordelingsmodel beoordeeld.

### 3.3.1 Ontwikkeling van de luistervaardigheid

Om de luistervaardigheid van leerlingen uit de peilingsonderzoeken van 2007 en 2017 met elkaar te vergelijken, kijken we naar de prestaties op twaalf ankeritems die in beide jaren door een deel van de leerlingen zijn gemaakt ( $n_{\text{totaal}}=1467$ ). Bij het vergelijken van prestaties tussen de twee peiljaren, is het belangrijk om te controleren voor factoren die van invloed kunnen zijn op de luistervaardigheid en die tussen peiljaren verschillen. Een factor die bleek te verschillen tussen peiljaren is het percentage leerlingen dat thuis een andere taal spreekt dan het Nederlands<sup>22</sup>. Ging het in 2007 nog om 8 procent van de groep leerlingen die de ankeritems maakte, in 2017 was dat 13 procent. Ook als we hiermee rekening houden, zien we dat leerlingen in 2007 beter presteerden op de ankeritems luistervaardigheid dan leerlingen in 2017. In 2007 werd gemiddeld 71 procent van de items goed gemaakt, terwijl in 2017 64 procent van de vragen correct werd beantwoord<sup>23</sup>.

### 3.3.2 Ontwikkeling van de spreekvaardigheid

Om de spreekvaardigheid van leerlingen uit de peilingsonderzoeken van 2010 en 2017 met elkaar te kunnen vergelijken, is de ankertoets spreken uit 2010 afgenomen bij een representatieve<sup>24</sup> groep leerlingen uit het huidige peilingsonderzoek ( $n=195$ ). Deze ankertoets bestond uit een beeldverhalentaak. Daarin moesten leerlingen een verhaal vertellen op basis van een serie afbeeldingen die ze te zien kregen. Dit moesten ze in totaal drie keer doen. Een van de afbeeldingenseries beeldde bijvoorbeeld een ruzie op het schoolplein uit, die uiteindelijk wordt gesust door de leerkracht.

Leerlingen werden beoordeeld aan de hand van het beoordelingsmodel uit 2010. Dit model bevatte de beoordelingsaspecten:

- Inhoud
- Woordkeuze
- Structuur en samenhang:
  - Naamgeving van de hoofdpersoon;
  - Metacognitieve en metalinguïstische werkwoorden (M-werkwoorden);
  - Aangeven van de setting;
  - Causale verbindingswoorden;
- Vloeiendheid ofwel aarzelend spreken;
- Verstaanbaarheid;
- Globaal kwaliteitsoordeel.

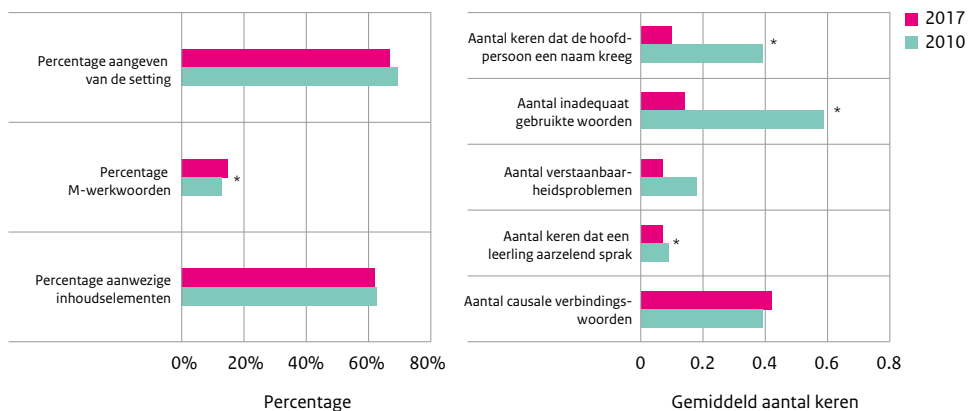
<sup>22</sup> Hier gedefinieerd als een andere taal dan het Standaardnederlands, een Nederlands dialect of een streektaal (bijvoorbeeld Fries).

<sup>23</sup> Hierbij moet opgemerkt worden dat de ankeritems gebaseerd waren op twee audiofragmenten: een fragment over Beau van Erven Dorens en een fragment over @Home internet. Items met vragen die horen bij het eerste fragment zijn in 2017 aanzienlijk slechter gemaakt dan in 2007 (proportie beheersing 0,63 versus 0,75). De scores op items die horen bij het tweede fragment verschilden echter niet veel van elkaar (proportie beheersing 0,68 en 0,66 in respectievelijk 2007 en 2017). Het verschil in prestaties op de items die horen bij de twee fragmenten is daarom misschien inhoudelijk te duiden.

<sup>24</sup> Deze groep leerlingen bleek voor wat betreft de resultaten behaald op het onderdeel spreekvaardigheid, representatief voor de gehele populatie, zie Van Langen et al., 2017.

Om de ontwikkeling van de spreekvaardigheid tussen 2010 en 2017 te onderzoeken, zijn de scores op de beoordeelde aspecten met elkaar vergeleken. Uit deze vergelijkingen komt naar voren dat beide groepen leerlingen nauwelijks van elkaar verschillen. De spreekvaardigheid van de leerlingen in groep 8 van het basisonderwijs is niet noemenswaardig verbeterd of verslechterd tussen 2010 en 2017. Verschillen zijn slechts zichtbaar op detailniveau (zie figuur 3.3.2a).

Figuur 3.3.2a – Gemiddelde scores (in percentage en aantal keer) op beoordeelde aspecten van het onderdeel spreekvaardigheid in de peiling van 2010 en 2017 ( $n_{2010}=204$  en  $n_{2017}=195$ ). Significante verschillen zijn aangegeven met een \*



Zo is het aantal inadequaet gebruikte woorden afgenomen tussen 2010 en 2017 (van gemiddeld 0,59 naar gemiddeld 0,14). Ook is het aantal keren dat een leerling aarzelend sprak (gemiddeld 0,09 keer in 2010 en 0,07 keer in 2017) iets minder geworden. Het aantal keren dat de hoofdpersoon in het verhaal een naam kreeg, wat bijdraagt aan de structuur en samenhang van het verhaal, is iets afgenomen (gemiddeld 0,39 keer in 2010 en 0,10 keer in 2017). Ten slotte is het gebruik van het aantal metacognitieve en metalinguïstische werkwoorden (“M-werkwoorden”) licht gestegen. M-werkwoorden zijn woorden die een innerlijke staat of een talige handeling weergeven, zoals ‘denken’, ‘voelen’ of ‘zeggen’ en ‘roepen’. Deze werkwoorden geven structuur aan een verhaal, doordat ze inzicht geven in de beweegredenen van de hoofdpersonen, of hun reacties op gebeurtenissen in het verhaal. Van de 25 van tevoren gespecificeerde M-werkwoorden gebruikte de gemiddelde leerling er in 2007 13. In 2017 waren dat er 15 (in de 3 beeldverhalen tezamen).

### TRENDS IN HET KORT

Leerlingen presteerden in 2007 beter op de ankeritems luistervaardigheid dan leerlingen in 2017. In 2007 werd gemiddeld 71 procent van de items goed gemaakt, terwijl in 2017 64 procent van de vragen correct werd beantwoord. Het blijkt dat de ankeritems van alleen 1 van de 2 audiofragmenten het verschil tussen 2007 en 2017 voor hun rekening nemen.

De spreekvaardigheid van de leerlingen in groep 8 van het basisonderwijs is niet noemenswaardig verbeterd of verslechterd tussen 2010 en 2017. Op detailniveau zijn kleine verschillen met 2007 zichtbaar, zowel in positieve als in negatieve zin.