QUICK MEASURES OF VOCABULARY:
DIVERSITY, WORD LEVEL COMPLEXITY AND PRODUCTIVITY

ASSESSING VOCABULARY

Students use of vocabulary can be measured in a number of different ways. Analysing students' language samples (e.g. a narrative generation, personal recount, persuasive argument or general conversation), can be a useful tool for tracking growth in students' language.

WORD COUNTS TOOL

OVERVIEW & USE

WordCountTools.com is a website can be used to quantify aspects of students' vocabulary use, including the:

- amount of words students use
- diversity of words students use
- length and complexity or words students use

Changes in these aspects generally reflect changes in students' overall expressive vocabulary.

How to use the tool:

- Open the website https://wordcounttools.com/
- In the big white box, enter a students' transcribed oral language (or written) sample
- On the right hand side of the web page, a number of statistics will be displayed – use these to analyse the students' vocabulary use (see below for details). These statistics can be used to track progress.

*Note: the **context** in which a language sample was collected will influence students' vocabulary use. Keep this in mind when analysing data.*
*Also **ensure that words are transcribed consistently** (e.g. 'um' always written as 'um', not 'ummm', 'umm' etc.) for accurate measures*

MOST USEFUL MEASURES & INFORMATION

Category	Measure	What it Is	Use. *May show changes in:*
Basic Word Count Statistics	**Words**	Total number of words in sample.	**Language production volume.**
Extra Word Count Statistics	**Unique Words**	Total unique or different words. *Number of words & percentage of total words*	**Language diversity.** An increase in the number of unique words used may reflect a wider vocabulary.
	Syllables per word	Average number of syllables per word.	**Basic word level language complexity.** *Increased average syllables per word may correlate with an increased application of affixes and use of more 'complex' vocabulary.* ***Note, not all complex words are long words. Vocabulary complexity cannot solely be measured by looking at word length.***
Top Keyword Density	**-**	List of most the used single words, 2 and 3 word combinations	This data may be useful for identifying specific words which the student uses frequently, in addition to word combinations used frequently e.g. using 'big bad' as descriptors often.

ADDITIONAL, LESS RELIABLE MEASURES:

Category	Measure	What it Is	Use. *May show changes in:*
Extra Word Count Statistics	**Monosyllabic Words (1 syllable)**	Total number of 1 syllable words (includes repeated words).	**Word level language production complexity.** *Increased use of polysyllable words may correlate with an increased application of affixes and use of more difficult vocabulary.*
	Polysyllabic Words (≥3 syllables)	Total number of 3+ syllable words (includes repeated words). Note, 2 syllable words are excluded from counts	
	Difficult Words	Number of words which are not on the website's list of '3000 familiar simple words'. ***CAUTION: This list is from research in the 1940s**	

* **CAUTION: The WordCountTools.com's difficult words list is the 'Dale-Chall 3,000 Word List for Readability'.** This is a list of just under 3000 familiar words developed in the 1940s. This list was drawn from words that were 'understood' when reading by at least 80 percent of the children in Grade 4 tested.

For more information see http://www.rfp-templates.com/research-articles/dale-chall-3000-simple-word-list#The_Dale-Chall_Word_List **Note, this list is different from other lists of common words in English – some researched and others not** e.g. the Oxford 3000 https://www.oxfordlearnersdictionaries.com/wordlist/english/oxford3000/ and the Longman Communication 3000 https://www.lextutor.ca/freq/lists_download/longman_3000_list.pdf.

CALCULATING THE NUMBER OF NEW WORDS DEMONSTRATED ACROSS 2 TIME POINTS

Use the NEMLDC Outreach Vocabulary Measures Datasheet (Number of New Words tab or Percentage of New Words tab), in conjunction with https://wordcounttools.com/

For ease of use it is suggested that student transcripts are entered into the NELMDC transcript template which can be downloaded here or accessed through the NEMLDC Outreach Service.
In this template both students' time 1 & 2 transcripts are entered into the same document, with one document per student. Student & assessment details should be entered into the headers of the document.

If using the transcript template:

1. Enter students' transcripts into the relevant pages of the transcript template (time 1 & time 2). Make sure to only enter what the student has said and write words/sounds consistently.

1. Copy & paste students' time 1 narrative transcripts into https://wordcounttools.com/ and look at the 'Extra Word Count Statistics' --> 'Unique Words' (number or %). Enter this number into the relevant 'Time 1' column.

2. Copy & paste everything in the students' document (aka both time 1 and time 2 transcripts) into the website.

Now look at the 'Extra Word Count Statistics' --> 'Unique Words' (number or %). Enter this number into the relevant 'Time 1 & Time 2 Combined' column. This shows the percentage / how many new unique words the student demonstrated at their second data point as compared to their first data point.

If not using the transcript template:

Note: For this analysis, make sure to only enter what the student has said and to write words/sounds consistently.
1. Enter a students' time 1 narrative transcript into https://wordcounttools.com/ and look at the 'Extra Word Count Statistics' --> 'Unique Words' (number or %). Enter this number into the relevant 'Time 1' column.

2. Leave that student's initial transcript in the website. Enter their time 2 transcript into the same box, so that you have the student's first and final transcripts in one box in https://wordcounttools.com/.

Now look at the 'Extra Word Count Statistics' --> 'Unique Words' (number or %). Enter this number into the relevant 'Time 1 & Time 2 Combined' column. This shows the percentage / how many new unique words the student demonstrated at their second data point as compared to their first data point.