Shaping innovative services: Reflecting on current and future practice

Language sample analysis

A powerful tool in the school setting

Samuel Calder, Cindy Stirling, Laura Glisson, Alannah Goerke, Tina Kilpatrick, Lauren Koch, Anna Taylor, Robert Wells and Mary Claessen

Language sample analysis is a useful method of evaluating children’s language performance. Computer-aided systems such as Systematic Analysis of Language Transcription (SALT) can serve to alleviate constraints clinicians face when analysing language samples to inform clinical decision-making. This article describes an initiative undertaken by a team of speech-language pathologists in a school context to enhance the efficiency and comprehensiveness of analysis of a narrative retell task in a sample of 131 children with developmental language disorder, using SALT. We report on the practicality of using SALT in this school context, and reflect on our experiences using the tool. We conclude that SALT is a valuable, evidence-based tool that enhances intervention planning and outcome measurement within the school-based clinical setting, and offers insights into future directions involving the use of systematic analysis of language transcripts within teams.

Demonstrating the effectiveness of services is challenging for all speech-language pathologists (SLPs). This paper reports on the process of systematic language sample analysis adopted by a team of SLPs employed in a Language Development Centre (LDC), a school for children with developmental language disorder (DLD). Intervention is provided at a classroom level in this setting; however, measuring children’s individual progress in addition to cohort-level outcomes is particularly important as each child’s placement within the specialist language centre is reviewed every year. As of 2017, the centre caters for approximately 260 students, with 23 teachers and 15 education assistants to provide classroom level intervention. A team of five SLPs operate within a responsiveness to intervention model (Gillam & Justice, 2010), providing direct specialised support to students at the whole class (Tier 1), small group (Tier 2) or individual level (Tier 3), or through consultation with educators in the centre. Given the large number of students with language support needs, SLPs at the centre must use time and resources efficiently to manage large caseloads, establish baseline language performance, plan and implement intervention, and demonstrate effectiveness of intervention. Dynamic data collection and analysis inform whether students can be discharged to mainstream schooling or whether their needs are best addressed at the LDC, and therefore as clinicians we regularly reflect on ways to improve the efficiency and effectiveness of our practices.

Tools for evaluating language performance

In order to establish baseline performance, SLPs can select from a number of tools available to assess language. Norm-referenced tests allow SLPs to compare children with age-matched peers in order to identify the presence of language disorders, whereas criterion-referenced tools measure a child’s performance of a particular linguistic skill in reference to a prior criterion of success (Paul & Norbury, 2012). Though norm-referenced assessments are useful for diagnosis, they are often limited in their capacity to measure change and lack cultural relevance for certain populations (Danahy Ebert & Scott 2014; Shipley & McAfee 2009). Therefore, one must also consider use of criterion-referenced tools such as language sample analysis (LSA). LSA supports evaluation of a child’s language performance in a naturalistic manner. LSA thus enables clinicians to collect and analyse data that represent linguistic performance across a range of real-life and structured communication tasks (Price, Hendricks, & Cook, 2010). It also allows SLPs to acquire data across a range of different genres and purposes that may be considered more ecologically valid (Dunn, Flax, Siwiński, & Aram, 1996). Furthermore, criterion-referenced tools such as LSA allow improvement in targeted skills to be evaluated in a dynamic way throughout intervention; in other words it is not as constrained as standardised norm-referenced tests regarding test-retest intervals (Paul & Norbury, 2012). Measuring oral language functioning by systematically analysing language samples for relevant criteria is often considered best-practice (Heilmann, Miller, Nockerts & Dunaway, 2010; Price et al., 2010).

Narrative language sampling

Within a school context, a range of genres may be sampled and analysed (Whitworth, Claessen, Leitão, & Webster, 2015); however, the importance of narrative performance is well recognised in the literature (Danahy Ebert & Scott, 2015). Language sample analysis is considered best-practice (Heilmann, Miller, Nockerts & Dunaway, 2010; Price et al., 2010) as it is not as constrained as standardised norm-referenced tests regarding test-retest intervals (Paul & Norbury, 2012). Measuring oral language functioning by systematically analysing language samples for relevant criteria is often considered best-practice (Heilmann, Miller, Nockerts & Dunaway, 2010; Price et al., 2010).
Narrative is considered a bridge between oral and literate language (Westby, 1985), and consequently, performance on narrative tasks is considered a strong predictor of academic success (Wellsman, Lewis, Freebairn, Avrich, Hansen, & Stein, 2011). Methods of analysing language performance through oral narrative are therefore useful for planning intervention to improve language-based academic outcomes, particularly at the classroom level (Spencer, Petersen, Slocum, & Allen, 2015). Narrative analysis offers information regarding language functioning at both the level of discourse (macrostructure) and the sentence and word level (microstructure). Such information enables SLPs to establish accurate and individualised intervention goals based on students’ needs (Spencer et al., 2015; Westerveld & Gillon 2008).

Although collection of a narrative sample is common practice for clinicians working with school-aged children, the time and effort required to complete a narrative analysis serves as a barrier to many SLPs (Pavelko et al., 2016). Pavelko, Owens, Ireland & Hahs-Vaughn, 2016; Westerveld & Claessen, 2014). Westerveld and Claessen (2014) reported that although 91% of Australian SLPs routinely collect language samples, only 37% undertake a detailed analysis. Reported barriers include time pressures and lack of training in using computer-assisted LSA. Similar findings were reported in a recent survey of 1,399 SLPs from the United States (Pavelko et al., 2016), suggesting that this is a widespread constraint. One method of implementing narrative sample analysis more efficiently is through the use of Systematic Analysis of Language Transcripts (SALT; Miller, Gillon, & Westerveld, 2015).

### Analysing language samples systematically

SALT (Miller et al., 2015) is a software tool that can be used to calculate microstructural language measures such as mean length of utterance (MLU) and number of different words (NDW). Such measures have been shown to correlate with norm-referenced test scores in identifying language disorder (Condouris, Meyer, & Tager-Flusberg, 2003). The software provides reference databases to compare performance to age- or grade-matched typical speakers on microstructure features, which may indicate disordered language performance compared to typically developing speakers (Norbury & Bishop, 2003). SALT can also be used to analyse a child’s use of macrostructural linguistic features, such as story grammar components in narrative retell tasks (Petersen, Gillam, & Gillam, 2008). Overall, the combination of narrative language sampling and analysis via SALT is an ecologically valid, dynamic and change-sensitive tool that utilises both norm-referenced and criterion-referenced processes to track language functioning.

Computer-aided systems like SALT enable SLPs to efficiently calculate a range of relevant measures which may inform diagnosis, treatment planning, and measurement of therapy effectiveness (Price et al., 2010). Results for an individual student or cohort may be compared to electronic databases, and individual scores may be compared across time to measure change on a range of performance criteria (Danahy Ebert & Scott, 2014; Petersen, Gillam, Spencer, & Gillam, 2010; Price et al., 2010). The use of such a tool has potential to alleviate some of the challenges faced by SLPs working with large caseloads of children with DLD and facilitate evidence-based practice. The introduction of new processes and clinical tools is challenging when working as part of a large team of SLPs within a school context and thus processes must be clearly documented to ensure consistency, clear communication and team alignment with the change.

### Purpose of the project

This paper describes the outcomes of a project designed to investigate the practicality of using SALT to systematically analyse the baseline narrative language samples of a large cohort of children with DLD from kindergarten to year 1 within an Australian specialised school context. As a large team of SLPs, we sought to pilot the use of SALT as a way to more efficiently analyse and use data to plan intervention and track progress, and to document the processes undertaken as well as our experiences with using the tool. We consider a number of factors associated with using SALT including elicitation and transcription of narratives, generation and application of codes, analysing baseline data at a cohort level, and the impact on classroom level intervention planning as well as team processes for managing service innovation and change. We also reflect on future directions for outcome measurement using SALT, with a particular emphasis on the clinical utility of systematic language analysis to inform discharge recommendations within a specialised school setting. Ethics approval was obtained from Curtin University (HREC2016-0047) and the Department of Education, Western Australia.

### Introducing SALT within a school context

#### The process for collecting narrative language samples

Narrative language sampling for 131 students with DLD was conducted at the end of 2015 to establish baselines across a range of language criteria and to set intervention goals for the following year. Although narrative sampling was already used as a standard part of assessment practice within the school, 2015 was the first year that the samples were analysed using SALT. Previously analysis occurred by hand using paper-based criterion-referenced rubrics such as those included in the Peter and the Cat narrative assessment tool (Allan & Leitão, 2003). Individual baseline data for each student, as opposed to cohort-level data, was our focus. To facilitate consistent elicitation of narratives, training and guidelines for narrative sampling procedures were provided to classroom teachers by SLPs. In some cases, this included SLPs modelling the elicitation of a narrative sample and providing explicit instruction on how to transcribe each sample verbatim (orthographic gloss). This was usually carried out 1:1 and took no more than 45 minutes. Extra support was provided if required.

All language samples were recorded using digital and analogue voice recorders and samples were transcribed verbatim by LDC classroom teachers. SLPs listened to the recorded samples and checked the teachers’ transcriptions, which were edited accordingly. Samples were then analysed by SLPs using SALT Research Version software (Miller et al., 2015). Language samples from pre-primary and year 1 students were elicited using Peter and the Cat (Allan & Leitão, 2003). For kindergarten students, Emma’s First Day narrative was used (West Coast LDC, unpublished assessment, see Appendix 1), as kindergarten-aged children fall below the recommended age range (5–9 years) for testing with Peter and the Cat. In both tasks, children were shown a wordless picture book as an accompanying story was read aloud to them. Children were then required to retell the story using the pictures as

---

<sup>From top to bottom: Alannah Goerke, Tina Kilpatrick, Lauren Koch and Anna Taylor</sup>
visual prompts. Both stories contain all key macrostructure elements (Stein & Glenn, 1979) and followed a similar structure to that of other standardised narrative discourse tests, such as the Bus Story (Renfrew, 1997).

SLP team procedures

A total of 7 SLPs and 15 teachers were involved in the process. Average years’ experience of the SLP team was 3.67 (SD 3.07; range 0.75–9.92) with a total of 6.2 full-time staff equivalent at the time of the project. In Western Australia, SLPs are trained in the use of SALT as part of university training; however, all team members participated in an extra 4-hour in-house training session to ensure consistency of segmentation and coding. SALT guidelines for segmenting and coding are freely available from http://www.saltsoftware.com/coursefiles/shared/Cunits.pdf and https://www.saltsoftware.com/salt/TranConvSummary.pdf, respectively. All transcripts were coded according to these guidelines. Samples were organised according to classroom and randomly allocated to SLPs depending on availability.

Language sample measures collected

Key standard measures were selected for evaluation and reporting to teachers as they are accepted measures of (a) expressive discourse (Total Number of Utterances, Number of Total Words), (b) syntax (MLU-morphemes), (c) semantics (NDW), and (d) verbal fluency (percentage of intelligible utterances and percentage of utterances with error) (Danahy Ebert & Scott, 2014).

SALT provides the option for clinicians to enter novel codes which prompt the software to identify specific features for analysis within samples. Novel codes were developed by the team to mirror narrative language elements measured by Monitoring Indicators of Scholarly Language (MISL; Gillam & Gillam, 2013), the Test of Narrative Language (TNL; Gillam & Pearson, 2004), and the Narrative Scoring Scheme (Heilmann et al., 2010), as well as those used in Petersen et al. (2010). For a full list of novel codes and corresponding narrative language elements see Table 1. Measures resulting from these codes were also reported to teachers.

Processes for analysing narrative language samples

Upon reflection, SLPs in the team reported requiring two to three hours, on average, to segment and code the first transcripts. However, as the team became familiar with the process, the average time each SLP spent per transcript reduced; in some cases falling to an average of 20–25 minutes per transcript which is comparable to the time previously spent analysing samples by hand. To ensure consistency of coding, SLPs worked together and provided support to each other where required. Where difficulties arose, team members discussed this and came to a consensus. These decisions were recorded in a log of “common issues” in order to ensure consistency. Most SLPs reported it easiest to segment and code several transcripts in one sitting, as this allowed clinicians to build “momentum”. Author SC checked 20% of samples from each class to ensure consistency in segmenting communication-units and coding (see SALT segmenting and transcription conventions above). While interrater

| Table 1. Project-specific macrostructure measures (adapted from Gillam & Gillam, 2013; Gillam & Pearson, 2004; Heilmann et al., 2010). |
| Macrostructure element and SALT code | Definitions |
| Orientation setting (place/time) [OS] | A reference to time and place relevant to the story (e.g., “One morning” or “walking home from school” excluding stereotypes like “Once upon a time”) |
| Orientation character [OC] | A reference to the agent of the story (e.g., “Emma” or “Peter”) |
| Additional character [CH] | A reference to any character that is not the agent within the story (e.g., “the man that helps Peter”) |
| Critical triangle – Initiating event [IE] | An event or problem that causes an emotional response from the character (e.g., “Mom left” or “Peter found a cat in a tree”) |
| – Internal response [IR] | Any reference to an emotional state (e.g., “Emma felt sad” or “Peter was worried”) |
| – Plan [P] | Reference to a cognitive verb indicating intention (e.g., “Emma decided to play with friends” or “Peter decided to climb up the tree”) |
| Actions [A] | Actions taken by the characters that are relevant to the story but not necessarily related to the initiating event (e.g., “Peter yelled for help”) |
| Emotion [E] | Emotions unrelated to the initiating event (e.g., “Peter was scared up the tree”) |
| Complication [COMP] | An event that stops the character from carrying out the plan related to the initiating event (e.g., “Peter is stuck in the tree”) |
| Solution/Resolution [S] | An event that resolves complication or initiating event (e.g., “Emma is picked up by her mum” or “The man helped Peter down from the tree”) |
| Consequence/tie up [C] | The outcome of the actions related to the initiating event (e.g., “Emma had a good first day at school” or “Peter’s mum told him to ask for help next time”) |
| Formulaic marker [FORM] | Standard utterances used to mark the boundaries of the narrative (e.g., “Once upon a time” or “The end”) |
| Character speech [SPEECH] | Any reference to character speech, both marked (e.g., Peter yelled, “Help!”) and unmarked (“Peter yelled for help”) |
databases, and individual or cohort scores may be used to plan evidence-based narrative intervention approaches (Spencer et al., 2015). Further, results can be compared across time to quantify change on a range of measures (Danahy Ebert & Scott, 2014).

Lessons learned and future directions
As a team of seven SLPs in a specialised school context, we explored an innovative way to more efficiently and systematically analyse cohort data to inform intervention planning. To achieve this we implemented systematic analysis of narrative samples using SALT. By the end of the project, all seven SLPs were confidently using SALT to check, code and analyse narrative language samples of a cohort of 131 preschool and school-aged children with DLD. The results of the analyses were used to establish baseline of children’s language functioning at a cohort level to guide classroom planning of narrative intervention. We considered this important because previous paper-based methods of analysis did not allow cohort-level data collation. The Rectangular Data File function in the software (also compatible with Microsoft Excel) allowed us to interpret and disseminate the information to school teaching staff in a clear and time efficient manner. Though the process of using SALT was initially time consuming and took longer than coding samples by hand, the team was able to obtain a greater depth of information using SALT across a range of macro- and microstructure narrative elements, which we feel has ultimately improved the quality of our baseline data collection and consequently the focus of our classroom level interventions, including small group and whole of class input.

Challenges and limitations
The project also facilitated reflection on assessment practices used at the LDC prior to and during the project,

![Figure 1. LDC percentage occurrence of macrostructure elements in pre-primary](image)

reliability was not calculated statistically, disagreement was minimal, likely as a result of the rigorous training process and collaborative coding of data.

The following sections discuss data that were used to support classroom planning of Tier 1 (whole class) intervention. In addition, the practical benefits and difficulties of using LSA in a school context are summarised.

Using narrative language sample measures to inform intervention planning
In order to inform both classroom level intervention goals and individual goals, percentage occurrence of narrative components were calculated for each year group. For example, 52% of pre-primary children did not use the macrostructure element “plan” (see Figure 1). The “plan” is an expansion of the traditional macrostructure elements (Stein & Glenn, 1979) linked to the “initiating event”. The element describes character’s plans to carry out actions in the story. This literary device is thought to support students to develop: comprehension of feelings; theory of mind; problem-solving and conflict resolution, and; the ability to plan for conversational interaction, among other important classroom-based skills. This was therefore selected as an intervention target for the class.

Importantly, the electronic aspect of SALT allowed for the collation of these kinds of data at the cohort level with ease using the “Rectangular Data File” function specific to the SALT Research Version. Previously, our team had been unable to focus our analysis and intervention planning at this level in an objective and systematic way. Attainment and use of these differentiated metrics is in-line with recommendations to implement the responsiveness to intervention model (Gillam & Justice, 2010), which is considered an evidence-based approach to supporting oral language development in an at-risk classroom. That is, results for individuals may be compared to electronic
and some challenges were identified. The main challenges were around elicitation procedures and transcribing procedures. These included some teachers not allowing students enough time to respond, and not labelling the transcripts. In order to overcome these challenges, the training protocol for teachers has been updated to include provision of direct modelling of narrative sampling techniques for all teachers, and ensuring instructions are in line with existing SALT elicitation procedures.

SALT was used to code a number of macro- and microstructure elements specific to the project. A further challenge was in the definitions developed for some of these SALT codes. Although definitions had been developed prior to the project commencing, some codes such as the “initiating event” required further discussion and resulted in further refinement to align with consensus in the literature (see Petersen et al., 2010).

Segmenting and coding transcripts was initially a confronting task as previously narrative sample analysis consisted of using paper-based resources to evaluate transcripts without systematically applying codes or transcription conventions. As the SLPs familiarised themselves with the SALT codes and coding conventions, they overcame the “fear” of using SALT. One team member in particular was quoted saying: “I had never used SALT since university because it seemed so daunting. With the support of the team, I can now see the value in using the software to systematically analyse language samples beyond the individual student.”

The SLP team who are now familiar with the codes and SALT conventions have since been able to segment and code transcripts with greater confidence and are committed to its use in future years.

Overcoming the challenges
It is anticipated that with ongoing practice, the SLPs in our team will continue to refine our competence and confidence with the process and in doing so optimise the viability of using SALT to efficiently and accurately measure language performance through narrative LSA. Future investigations may include calculating interrater reliability using intra-class correlations for 20% of the samples to determine the overall accuracy of using the tool. Further, it may be useful to systematically determine a time breakdown of the total time to record, transcribe, check, code and analyse all transcriptions to evaluate the efficiency of using such a tool in place of paper-based standardised tests. In addition, Pavelko et al. (2016) outline four best-practice principles to language sampling, including: (a) use of narrative sampling contexts; (b) obtaining samples with a length of at least 50 utterances, (c) recording and transcribing samples; and (d) selecting computer software for efficient analysis and interpretation of data. In our paper, three of the four principles have been achieved. By altering sampling contexts to obtain 50 utterances, all four principles could be employed, indicating that it is achievable to use LSA as an evidence-based approach to planning intervention for children with DLD in a school context.

Future directions
In future studies, we plan to use the information collected through systematic cohort-level LSA to evaluate class and individual level data in order to monitor and report on progress. For example, the first step will be to analyse the same macro- and microstructure measures one-year post initial language sampling using the process described above. Pre- and post-measures can be compared across individuals to determine if progress in targeted areas of narrative intervention have improved. Further, data can be collapsed using the Rectangular Data File function within SALT, and used to calculate descriptive statistics at a cohort level and thus to determine whether change across measures is consistent across year levels.

We recognise that arriving at a conclusion as to whether or not language intervention has been effective must account for confounding factors such as maturation and environmental/ history effects. It is acknowledged that pre-post comparisons in isolation are not especially robust for achieving this purpose. Nonetheless, the strength of this evaluation of progress can be tested (see Pring, 2005), and is clinically useful if considered in conjunction with other methods of evaluating effectiveness (see Ebbels, 2017). Therefore, future directions will include the continued use of SALT to evaluate the effectiveness of the LDC Tier 1 narrative oral language program at an individual and cohort level.

Conclusions
In summary, the SLP team at the LDC found SALT to be a valuable clinical tool that is transferable to the school context with some local adaptations. As with any new clinical practice tool or process, extra resources were required initially. Cooperation from the whole team and support from school administration were vital to the success of the project as were acknowledgement and acceptance of the need for initial investment of time and resources and an ongoing commitment to evaluation and reflection. This article demonstrates that using narrative language sampling and SALT within a school context is achievable, even with large numbers of students. It is an efficient and evidence-based approach to systematic analysis of data which has potential to enhance planning of intervention, comparison and review of language performance at both an individual and cohort level, and ultimately the efficacy of speech language pathology interventions within a school context.

References


Samuel Calder is a speech pathologist and support officer, speech & language at North East Metropolitan Language Development Centre. Cindy Stirling is a senior speech pathologist and support officer, speech & language at North East Metropolitan Language Development Centre. Laura Gilsson is outreach coordinator and a support officer, speech & language at North East Metropolitan Language Development Centre. Alannah Goerke is a support officer, speech & language at North East Metropolitan Language Development Centre. Tina Kilpatrick is a support officer, speech & language at North East Metropolitan Language Development Centre. Lauren Koch is a speech pathologist at North East Metropolitan Language Development Centre. Anna Taylor is a senior speech pathologist at North East Metropolitan Language Development Centre. Robert Wells is a speech pathologist and support officer, speech & language at North East Metropolitan Language Development Centre. Dr Mary Claessen is a speech pathologist and senior lecturer at Curtin University.

Correspondence to:
Samuel Calder
Department of Education, Western Australia
phone: (08) 0275 5511
e-mail: samuel.calder@education.wa.edu.au