LSHSS

Letter to the Editor

Is Putting SUGAR (Sampling Utterances of Grammatical Analysis Revised) Into Language Sample Analysis a Good Thing? A Response to Pavelko and Owens (2017)

Ling-Yu Guo,[a,b] Sarita Eisenberg,[c] Nan Bernstein Ratner,[d] and Brian MacWhinney[e]

Purpose: In this letter, the authors respond to Pavelko and Owens' (2017) newly advanced set of procedures for language sample analysis: Sampling Utterances and Grammatical Analysis Revised (SUGAR).
Method: The authors contrast some of the new guidelines for transcription, morpheme segmentation, and language sample elicitation in SUGAR with traditional conventions for language sample analysis (LSA). They address the potential impact of the new guidelines on some of the target measures in SUGAR—mean length of utterances in morphemes (MLU_m), words per sentence (WPS), and clauses per sentence (CPS)—and provide their suggestions.
Results: Inclusion of partially intelligible utterances in SUGAR may over- or underestimate children's MLU_m and reduce the reliability of computing WPS. Counting derivational morphemes and the component morphemes of catenatives (e.g., *gonna*) may result in overestimation of children's morphosyntactic skills.
Conclusion: Further data are needed to determine whether MLU_m including derivational morphemes and the component morphemes of catenatives is a better measure of children's morphosyntactic skills than MLU_m excluding those morphemes. Pending such data, the authors recommend maintaining traditional LSA conventions and measures. Furthermore, free, fast automated utilities already exist that reduce barriers for clinicians to conduct informative, in-depth LSA.

In evaluating children's language skills, clinicians must use a variety of technically sound assessment tools and strategies (Individuals with Disabilities Education Act, 2004). In addition to standardized tests, language sample analysis (LSA) has been considered an effective tool for assessing children's language skills (Bernstein Ratner & MacWhinney, 2016; Oetting et al., 2010; Paul & Norbury, 2012). Despite its effectiveness, a recent national survey indicated that only two thirds of the speech-language pathologists surveyed used LSA during the evaluation process (Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016).

Pavelko and Owens (2017) discussed several barriers to the clinical use of LSA, including, but not limited to, (a) the time-consuming nature of LSA; (b) the lack of clear and consistent guidelines for collecting, transcribing, and analyzing language samples; and (c) the lack of appropriate normative data for LSA measures. To address these issues, Pavelko and Owens developed Sampling Utterances and Grammatical Analysis Revised (SUGAR) to enable clinicians to conduct LSA efficiently. SUGAR uses readily available technology (e.g., smartphones, word processors) to transcribe the language samples and compute the measures. Using SUGAR, Pavelko and Owens reported that it took less than half an hour to collect, transcribe, and analyze language samples of 50 utterances (excluding utterances from the child's conversational partner). They provided rules for calculating each of four measures—mean length of utterances in morphemes (MLU_m), total number of words, words per sentence (WPS), and clauses per sentence (CPS)—and provided comparison reference data for

[a]Department of Communicative Disorders and Sciences, University at Buffalo, NY
[b]Department of Audiology and Speech-Language Pathology, Asia University, Taichung, Taiwan
[c]Department of Communication Sciences and Disorders, Montclair State University, Bloomfield, NJ
[d]Department of Hearing and Speech Sciences, University of Maryland, College Park
[e]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA
Correspondence to Ling-Yu Guo: lingyugu@buffalo.edu

Editor-in-Chief: Shelley Gray
Editor: Kerry Ebert
Received August 16, 2017
Revision received November 9, 2017
Accepted November 22, 2017
https://doi.org/10.1044/2018_LSHSS-17-0084

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

children between 3;0 (years;months) and 7;11 on 6-month or 1-year intervals based on their suggested transcription and analysis guidelines.

We applaud the development of a novel, heuristic method for clinicians to conduct LSA. It is important to note, however, that Pavelko and Owens (2017) fundamentally changed the way in which LSA measures are to be calculated, especially MLU_m. These changes are substantially different from the conventions established by Brown (1973), whose conventions are widely accepted in the field (e.g., MacWhinney, 2017; Miller, Andriacchi, & Nockerts, 2016; Paul & Norbury, 2012; Rice et al., 2010). In this response, we highlight some changes suggested by Pavelko and Owens, discuss their potential impact on LSA measures, and provide our suggestions to those changes whenever appropriate. We will first address specific issues about transcription and morpheme segmentation before addressing some more general issues about eliciting and analyzing language samples. Note that SUGAR also provides new guidelines for computing clausal density (i.e., CPS), which are different from the current way of computing clausal density (e.g., Heilmann & Malone, 2014; Nippold et al., 2014). To maintain the focus of this response letter, however, we will not discuss the guidelines for computing CPS (e.g., what counts as a sentence) in SUGAR.

Transcription and Morpheme Segmentation

Defining Unintelligible Utterances

In conventional LSA, only fully intelligible utterances are included; utterances with even one unintelligible word would be excluded from the analyses (Brown, 1973). In contrast, Pavelko and Owens (2017) defined unintelligible utterances as those with "three or more unintelligible words" (p. 204). This means that utterances with one or two unintelligible words would still be included in the SUGAR analyses. Pavelko and Owens suggested that this "three unintelligible words" rule is similar to the conventions of Systematic Analysis of Language Transcripts (SALT; Miller et al., 2016). However, to the best of our knowledge, SALT differentiates between unintelligible and fully intelligible utterances in their analyses. Utterances containing even one unintelligible segment are excluded from most calculations, including MLU_m, which was confirmed by Andriacchi (personal communication, July 31, 2017). In addition, the authors referenced by Pavelko and Owens in their Table 1 (Pavelko & Owens, 2017, p. 198) have all agreed upon excluding utterances with unintelligible segments because one cannot reliably compute MLU_m, WPS, or other values without knowing what the missing information really was.

There are two potential issues that could arise from this "three unintelligible words" rule. First, when an utterance has unintelligible segments, it could be difficult to determine the number of unintelligible words within those segments because the boundaries between unintelligible words may not always be clear. For example, an utterance with two unintelligible words could be mistakenly counted as one with three or more unintelligible words and vice versa. This could affect utterance inclusion and the counting of number of words in sentences. Second, including utterances with unintelligible segments could potentially underestimate a child's MLU_m. Consider the example in which a child says, "*He likes cookies,*" but the last two words are unintelligible. In SUGAR, this utterance would be transcribed as, "*He XX XX,*" and would be counted as having three morphemes. However, the child actually would have produced five morphemes even though the last two words were unintelligible.

Whether including utterances with one or two unintelligible words would significantly affect the computation of LSA measures still awaits empirical research. Until then, we recommend maintaining the convention that partially unintelligible utterances be excluded from quantitative LSA measures such as MLU_m because this is consistent with current clinical practice (Paul & Norbury, 2012).

Counting Derivational Morphemes

Bound morphemes can be divided into two types. *Inflectional morphemes* add a grammatical property to a word—such as noun plurality (e.g., *cat*; *cats*) or verb tense (e.g., *walk*; *walked*)—but do not change the meaning or syntactic class of the base word. *Derivational morphemes* create new words that have different meanings (e.g. *pure*; *impure*) and/or a different syntactic class (e.g., *teach*; *teacher*) from the base word. Brown (1973) excluded derivational morphemes in the calculation of MLU_m. He also excluded the *–ing* and *–ed* forms when they were used to create deverbal adjectives used as prenominal modifiers (e.g., *that is an interesting book*) and as predicate adjectives (e.g., *I am tired*). This is because those *–ing* and *–ed* forms, by definition, are derivational morphemes given that they change the syntactic class of the base word. In contrast, Pavelko and Owens (2017) included a number of derivational morphemes (including the derivational uses of *–ing* and *–ed*) in their computation of MLU_m. The rationale behind this decision was that children older than age 4 years used such derivational morphemes.

We recognize that including derivational morphemes is a novel, alternative way for computing MLU_m; we would also like to restate the reasons why derivational morphemes have not been included in the computation of MLU_m in traditional LSA. First, children may learn a derived word (e.g., *beautiful, interesting*) before the base word (e.g., *beauty, interest*). For example, a check of the Brown corpus in the Child Language Data Exchange System (CHILDES; MacWhinney, 2000) shows five early uses of *interesting* (all from Adam at 2;10, 2;11, 3;1, and 3;2) and no uses of *interest.* It seems unlikely that children would add the derivational morpheme (e.g., *–ful,* derivational *–ing*) to the base word to form the derived word in those cases.

Second, there is ample psycholinguistic evidence that base words and derived words (e.g., *beauty, beautiful*) are stored as separate lexical entries (e.g., Levelt, 1989; Libben, Jarema, Derwing, Riccardi, & Perlak, 2016). That is, a speaker stores each derived word as a whole in the mental lexicon although he or she may recognize the link between base words and derived words (i.e., morphological awareness). Thus, when speakers produce a derived word in an utterance (e.g., *beautiful*), it does not necessarily mean that they add the derivational morpheme to the base word during the language production process. Note that this is different than the process of producing words with inflectional morphemes, which is generative and additive rather than holistic.

Third, derived words vary considerably in the transparency of the base words. Semantically nontransparent base words are treated as monomorphemic words (Marslen-Wilson, Tyler, Waksler, & Older, 1994). Consider the two examples provided by Pavelko and Owens for the morpheme *–sion*: *discussion* and *mission* (Pavelko & Owens, 2017, p. 203, Table 3). The base word *discuss* (with the meaning "to talk about something") is a common word in daily usage so that *discussion* is readily separated into its component morphemes. However, the base word *miss* (with the meaning "to send") is an obsolete word that is no longer used (*Merriam-Webster's Online Dictionary,* 2015). Even adults, therefore, may not recognize the derivation of the word *mission*. This could potentially affect interrater reliability for segmenting derived words because different people may have different understandings of morphological compositions for words.

Last, and perhaps most crucially, derivation is a word-formation process, not a grammatical encoding process (Bock & Levelt, 1994). Therefore, derivation reflects a child's lexical skills, not grammatical skills. Given that MLU_m is considered a measure of children's "morphosyntactic complexity" (Pavelko & Owens, 2017, p. 198), it has yet to be determined whether counting derivational morphemes in the MLU_m computation would yield a valid measure for children's morphosyntactic skills, and if so, whether it is a better measure of children's morphosyntactic skills than the traditional MLU_m computation. These questions are important for clinicians because counting derivational morphemes requires additional time for morpheme segmentation. One way to answer these questions would be to compute the MLU_m values in the new way (i.e., counting derivational morphemes) and in the traditional way and then compare the magnitude of correlations between each of the MLU values and a reference standard of children's morphosyntactic skills (e.g., Structured Photographic Expressive Language Test–Preschool, Second Edition; Dawson et al., 2005). This piece of empirical evidence would help clinicians determine whether there is indeed added value in counting derivational morphemes in the MLU_m computation.

From a psycholinguistic perspective, we recommend maintaining the convention that derivational morphemes not be counted as separate morphemes for calculating MLU_m unless the empirical evidence suggests otherwise. Even if empirical evidence does support the inclusion of derivational morphemes, an ensuing issue is whether clinicians would need to count the other derivational morphemes that are not listed in Pavelko and Owens (2017, Table 3). Pavelko and Owens included only a limited number of derivational morphemes in the SUGAR analysis based on their pilot study. A situation that clinicians will encounter is children using words with derivational morphemes that are not listed (e.g., *–ive, –ity* as in *activity*). To compare children's performance to the reference data in Pavelko and Owens, clinicians would need to count the word *activity* as one morpheme, instead of three morphemes (i.e., *act, –ive, –ity*), because the morphemes *–ive* and *–ity* are not listed in the current SUGAR analysis. However, this would underestimate children's ability in using derivational morphemes in the utterances. Thus, we suggest that Pavelko and Owens be more explicit regarding how clinicians should handle words with unlisted derivational morphemes.

Counting Inflectional Morphemes –er and –est

Pavelko and Owens (2017) included the comparative *–er* and the superlative *–est* among the morphemes counted in SUGAR. Although *–er* and the *–est* are inflectional morphemes (Biber, Johansson, Leech, Conrad, & Finegan, 1999), Brown (1973) did not count these morphemes because they are not obligatory. This means that it is a stylistic choice among speakers whether to use comparative and superlative forms rather than the uninflected adjective. For instance, when faced with three balls of varying sizes, a child might pick up the largest one and say "*I have the big one*" unless prompted by the clinician to make a comparison. As noted by Johnston (2001), discourse-based variability arising from differences in the content and context of interaction can affect the choices children make in formulating their utterances and have a major influence on MLU_m. In the absence of empirical evidence showing that including these morphemes results in a valid, more sensitive measure of MLU_m, we recommend maintaining the convention that *–er* and *–est* not be counted as separate morphemes for calculating MLU_m as we suggested for derivational morphemes.

Counting Catenatives

Brown (1973) counted catenatives (e.g., *gonna, wanna, gotta, hafta*) as single morphemes because children may have stored catenatives as holistic chunks. In contrast, Pavelko and Owens (2017) counted catenatives as multiple morphemes. For example, they counted *wanna* as two morphemes (i.e., *want to*) and *gonna* as three morphemes (i.e., *go –ing to*). The reason given by Pavelko and Owens for treating catenatives as multiple morphemes was that "typical school-age children know the individual morphemes" for the catenatives (p. 203).

We agree with Pavelko and Owens (2017) that children may figure out the component morphemes of catenatives at some point of language development. This fact was also noted decades ago by Limber (1973), who noticed a different pattern of emergence of catenatives and their component pieces when they are separated by intervening verbal materials (e.g., "*I wanna go*" vs. "*I want John to go*"). Limber's conclusion—not seriously contested in the developmental literature—was that catenatives may indeed be lexicalized in young children's speech, which argues against inflating their value in LSA. There is evidence that adult speakers may store and activate catenatives holistically as chunks, instead of morpheme by morpheme, in language production (e.g., Bybee, 2006). Thus, it cannot be assumed that older children produce catenatives by first combining the component morphemes together and then pronouncing them in reduced forms. This assumption confounds the process of sentence production with morphological awareness about the component morphemes of catenatives.

To compare children's performance to the reference data in Pavelko and Owens (2017), clinicians are required to follow the novel rule by counting catenatives as multiple morphemes. However, clinicians may need to be aware that counting catenatives as multiple morphemes could potentially inflate children's MLU_m values and hence overestimate their morphosyntactic skills. Further research is needed to examine this issue.

Transcribing Interactant Utterances

In SUGAR, clinicians transcribe only the child's utterances but not the adult conversational partner's utterances. We recognize that this decision reduces transcription time. However, clinicians need to be aware of two caveats about not transcribing the conversational partner's utterances. First, it could be difficult to determine whether a child's utterance is an imitation of a previous utterance produced by the conversational partner. Traditionally, imitated utterances are excluded from analysis because they are not independently generated by the child (Brown, 1973). Without further notations, those imitated utterances could be mistakenly included in the computation.

Second, it could be difficult to determine whether subjectless utterances should be counted as a sentence for the analyses of words per sentence and clauses per sentence. Note that in SUGAR, elliptical utterances are counted as sentences "if some portion of the verb phrase is present" (Pavelko & Owens, 2017, Appendix C, p. 214). For instance, in response to the examiner's question "*What are you doing?*" the child could answer "*Running*." The utterance "*Running*" is an ellipsis in this context and would be counted as a sentence in the measure of words per sentence and clauses per sentence. However, in a different context, the child could label an ongoing action as "*running*" to refer to a horse that is running. In this case, this utterance "*running*" is not an ellipsis and would not be counted as a sentence. Without further notations, the reliability of identifying appropriate sentences for analysis could be compromised.

It may not be necessary to transcribe all utterances from the conversational partner for calculating child LSA measures such as MLU_m. We, however, recommend that the conversational partner's utterances be transcribed at least when they are necessary for determining utterance inclusion. Finally, it should be recognized that the well-discussed ecological validity of LSA, as opposed to standardized test performance (Hewitt, Hammer, Yont, & Tomblin, 2005), is diminished when the record of the conversation is reduced to one speaker's contributions. It decontextualizes the record of the child's performance and may limit subsequent analysis of the appropriateness of the child's conversational turns.

Language Sample Elicitation Contexts

Pavelko and Owens (2017) provided verbal techniques for clinicians to elicit conversational language samples (Appendix A, p. 212). These techniques indeed reflect conventional guidelines for maximizing sample representativeness (e.g., Retherford, 2007). Conversational language samples can be collected using different sampling contexts (e.g., free play, interview, picture description), which may affect the outcome of LSA measures. To give one example, Southwood and Russell (2004) reported differences in MLU_m and proportion of complex sentences between play samples (i.e., samples elicited while children played with toys) and interview samples (i.e., samples elicited by introducing topics and asking follow-up questions and comments). Thus, when a child's performance is compared to reference or normative data, it is important to replicate the specific procedures used for gathering that comparison data (Eisenberg, Fersko, & Lundgren, 2001; McCauley & Swisher, 1984). Although Pavelko and Owens provided clear guidelines for the verbal techniques to elicit conversational samples, they were not as explicit in describing which type of sampling contexts were used to generate the reference data or whether the sampling contexts varied with age in the reference data. Clinicians may, thus, not be able to determine whether the conversational samples they collect are comparable to those in the SUGAR reference data, which could lead to uncertainty in making clinical decisions. Therefore, we suggest that Pavelko and Owens provide specific information regarding sampling contexts by age to allow legitimate comparisons against the SUGAR reference data.

Concluding Thoughts

Pavelko and Owens (2017) developed SUGAR to address several barriers that limit clinical use of LSA: excessive time, inconsistent guidelines to conduct LSA, and lack of normative data. We agree with Pavelko and Owens that LSA provides valuable information, and we would also like to see increased use of LSA by clinicians. SUGAR analysis does significantly reduce the amount of time for LSA. Recall that SUGAR reduces transcribing time by skipping the conversational partner's utterances and child mazes (e.g., filled pauses like *um, uh,* repetition, revisions) and by limiting the analysis to four basic LSA measures (e.g., MLU_m). Traditional LSA could take longer time for coding when clinicians plan to compute additional grammatical measures other than the basic measures for making clinical decisions, such as percent tense usage and percent grammatical utterances (Eisenberg & Guo, 2016). Time demands for traditional LSA could be significantly reduced if clinicians skip the conversational partner's utterances and child mazes and also limit the analysis to basic LSA measures (e.g., MLU_m, total number of words, number of different words, type-token ratio).

Rather than reducing the content and output of LSA to save time and effort, clinicians should be aware that recent computerized LSA programs, such as Computerized Language Analysis (CLAN; MacWhinney, 2017), can automatically perform grammatical (i.e., morphological, syntactic) coding (Bernstein Ratner & MacWhinney, 2016). This means that the clinician can simply transcribe child utterances orthographically as Pavelko and Owens (2017) suggested in Appendix B (p. 213) and can calculate MLU_m and a wide variety of other grammatical measures without having to first segment utterances and morphemes, as still required by SUGAR. Before concluding that SUGAR saves time, studies are needed that directly compare SUGAR and computerized programs like CLAN that allow for automated grammatical analysis.

Finally, Pavelko and Owens (2017) were concerned that clinicians are faced with contradictory information about how to compute MLU_m across normative databases. Although it is true that different LSA databases may use somewhat different protocols for language sample collection, transcription, and analysis, this is not necessarily problematic. Just as each standardized test has its own procedures for testing and scoring, each LSA database will have its own protocol. As long as the protocol is specified so that clinicians follow that protocol, they can reliably compare the child's data to the norms in that database.

What we like most about SUGAR is that it addresses a real problem in clinical assessment and provides an innovative, heuristic method to motivate clinicians to conduct LSA. Its aim is to be user-friendly and provide an alternative to traditional LSA conventions for documenting developmental changes of children's language skills; it also adds reference data for four outcome measures from 3;0 to 7;11. However, we would like to point out that traditional LSA conventions are time efficient as well if clinicians use an appropriate tool (e.g., CLAN). In addition, LSA is meant to go beyond flagging diagnostic concern and provide information that can be used in therapy planning. It has yet to be determined what information SUGAR can provide clinicians to use in discriminating typical from atypical performance and then creating useful and ecologically valid interventions.

LSA has evolved significantly over the past 30 years. In this digital era, computer software that automatically analyzes language samples and reduces the complexity of LSA (e.g., CLAN) is now readily and freely available. New LSA measures that are sensitive to developmental changes and language impairment have also been created, such as finite verb morphology composite (Bedore & Leonard, 1998), tense and agreement productivity scores (Hadley & Short, 2005), and percent grammatical utterances (Eisenberg & Guo, 2016). We hope that discussion of SUGAR will shine light on the value of LSA in clinical assessment and planning and spur development and validation of other sensitive LSA measures that exploit current technology and archival resources.

Acknowledgments

This work was supported by FluencyBank grants from the National Institute on Deafness and Other Communication Disorders (1R01DC015494) and National Science Foundation–Social, Behavioral, and Economic Sciences (SMA-1539010) to Nan Bernstein-Ratner and Brian MacWhinney, and the National Institutes of Health CHILDES grant (R01-HD23998) to Brian MacWhinney.

References

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41,* 1185–1192. https://doi.org/10.1044/jslhr.4105.1185

Bernstein Ratner, N., & MacWhinney, B. (2016). Your laptop to the rescue: Using the Child Language Data Exchange System archive and CLAN utilities to improve child language sample analysis. *Seminars in Speech and Language, 37,* 74–84.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, United Kingdom: Pearson Education Limited.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*(4), 711–733.

Dawson, J., Stout, C., Eyer, J., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test–Preschool 2 (SPELT-P2)*. DeKalb, IL: Janelle Publications.

Eisenberg, S., Fersko, T., & Lundgren, C. (2001). Use of MLU for identifying language impairment in preschool children: A review. *American Journal of Speech-Language Pathology, 10,* 323–342. https://doi.org/10.1044/1058-0360(2001/028)

Eisenberg, S., & Guo, L.-Y. (2016). Using language sample analysis in clinical practice: Measures of grammatical accuracy for identifying language impairment in preschool and school-aged children. *Seminars in Speech and Language, 37*(2), 106–116. https://doi.org/10.1055/s-0036-1580740

Hadley, P., & Short, H. (2005). The onset of tense marking in children at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research, 48,* 1344–1362. https://doi.org/10.1044/1092-4388(2005/094)

Heilmann, J., & Malone, T. O. (2014). The rules of the game: Properties of a database of expository language samples. *Language, Speech, and Hearing Services in Schools, 45,* 277–290. https://doi.org/10.1044/2014_LSHSS-13-0050

Hewitt, L. E., Hammer, C. S., Yont, K. M., & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38*(3), 197–213.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Johnston, J. R. (2001). An alternate MLU calculation: Magnitude and variability of effects. *Journal of Speech, Language, and Hearing Research, 44,* 156–164. https://doi.org/10.1044/1092-4388(2001/014)

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Libben, G., Jarema, G., Derwing, B., Riccardi, A., & Perlak, D. (2016). Seeking the *-ational* in *derivational* morphology. *Aphasiology, 30*(11), 1304–1324.

Limber, J. (1973). The genesis of complex sentences. In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 169–185). New York, NY: Academic Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). New York, NY: Psychology Press.

MacWhinney, B. (2017). *Tools for analyzing talk. Part 2: The CLAN program*. Pittsburgh, PA: Carnegie Mellon University. Retrieved from http://talkbank.org/manuals/CLAN.pdf

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101,* 3–33.

McCauley, R. J., & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49,* 34–42. https://doi.org/10.1044/jshd.4901.34

Merriam-Webster's Online Dictionary. (2015). Retrieved August 10, 2017, from http://www.merriam-webster.com/

Miller, J., Andriacchi, K., & Nockerts, A. (2016). *Assessing language production using SALT software: A clinician's guide to language sample analysis* (2nd ed.). Madison, WI: SALT Software.

Nippold, M. A., Frantz-Kaspar, M. W., Cramond, P. M., Kirk, C., Hayward-Mayhew, C., & MacKinnon, M. (2014). Conversational and narrative speaking in adolescents: Examining the use of complex syntax. *Journal of Speech, Language, and Hearing Research, 57,* 876–886.

Oetting, J. B., Newkirk, B. L., Hartfield, L. R., Wynn, C. G., Pruitt, S. L., & Garrity, A. W. (2010). Index of Productive Syntax for children who speak African American English. *Language, Speech, and Hearing Services in Schools, 41,* 328–339. https://doi.org/10.1044/0161-1461(2009/08-0077)

Paul, R., & Norbury, C. F. (2012). *Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating* (4th ed.). St. Louis, MO: Elsevier.

Pavelko, S. L., & Owens, J. R. E. (2017). Sampling Utterances and Grammatical Analysis Revised (SUGAR): New normative values for language sample analysis measures. *Language, Speech, and Hearing Services in Schools, 48,* 197–215. https://doi.org/10.1044/2017_LSHSS-17-0022

Pavelko, S. L., Owens, R. E., Jr., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools, 47,* 246–258. https://doi.org/10.1044/2016_LSHSS-15-0044

Retherford, K. (2007). *Guide to analysis of language transcripts* (3rd ed.). Austin, TX: Pro-Ed.

Rice, M. L., Smolik, F., Perpich, D., Thompson, T., Rytting, N., & Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research, 53,* 333–349. https://doi.org/10.1044/1092-4388(2009/08-0183)

Southwood, F., & Russell, A. F. (2004). Comparison of conversation, freeplay, and story generation as methods of language sample elicitation. *Journal of Speech, Language, and Hearing Research, 47,* 366–376. https://doi.org/10.1044/1092-4388(2004/030)