

Using an AS-unit Complexity Benchmark to Measure Beginner Learner Oral Production in Communicative Tasks

Jason MOSER

Abstract

In this paper I will discuss learner language complexity in oral tasks and how it has been conceptualized in task-based learning research. One of the key tools for measuring complexity in spoken discourse is the AS-unit (Analysis of Speech Unit). The AS-unit is a main clause and any attached subordinate clauses or sub-clausal units. I will first discuss the problems involved in codifying AS-units in data from communicative pair work tasks. I will demonstrate that subordination is not necessarily characteristic of communicative tasks, nor is it easy to identify in conversation, especially with beginner learners. In response I will argue that measuring learner productivity by words per AS-unit is an effective alternative. I will also demonstrate an AS-unit complexity benchmark based on AS-unit word count, in which units above a certain word count are deemed complex. The rationale for this benchmark will be discussed, and supported with examples from four beginner learners' data.

Introduction

Currently there is very little research that provides an extensive description of how AS-units are codified in spoken data including how to decide whether clauses are part of a larger AS-unit or constitute their own AS-unit. Foster et al. (2000) provide guidance, but the application of their rules is applied to a short transcript of elliptical language, which has very little subordination and coordination. In task-based research language complexity has mostly been measured through the amount of subordination occurring in narrative tasks. This is not surprising since subordination is characteristic of the discourse structure of narratives, in that most narratives involve the describing of relationships of intentionality and causality (see Brown 1989). Berman and Slobin (1994) write that the simplest narratives contain a temporal sequence of simple clauses. Narratives at their most complex involve 'a hierarchical layering of circumstances and happenings in syntactically packaged constructions...' (14). Ellis and Barkhuizen's (2005) work demonstrates that codifying AS-units in narrative tasks is fairly straightforward. In the next section I will discuss how much more problematic it is to do for communicative pair work tasks.

Lack of Subordination in Conversation

Conversation cannot be neatly described by traditional grammatical description. McCarthy (1998: 79) explains that informal spoken data is typically characterized by lack of ‘well-formed sentences’ with main and subordinate clauses. The reason for this problem relates to the nature of what is involved in informal spoken conversation. Online production in relation to conversation has three basic principles: keep talking, limited planning ahead, and qualification of what has been said (Biber et al. 1999: 1067). These realities limit the degree of linguistic complexity possible, especially with L2 learners. What is normally visible in spoken discourse according to McCarthy are turns that have incomplete chunks of language including clauses that look like subordinate clauses, but are in fact not attached to a main clause. Carter & McCarthy (2006: 170) write:

The needs of real-time communication do not allow the speaker time to construct over-elaborate patterns of main and subordinate clauses. Much more common are sequences of clauses linked by coordinating conjunctions . . . or by simple subordinating conjunctions such as because (frequently contracted to cos) and so, which often function more like coordinating rather than subordinating conjunctions.

Carter & McCarthy (2006: 557) add that this multiple coordination is more commonly comprised of coordinators like *and* and *but*. Subordination is considered to be a basic form of integration, but nonetheless the ability to identify it is problematic when learners produce clauses in a continuous chain. Schleppegrell (1992: 119) on this point writes that identifying a main clause, and its accompanying subordinate clause is extremely difficult because the subordinate clause can in an expansive utterance become the ‘ideational core’ of the expansion, which is then supported by other clauses that follow. This problem becomes very acute when you are dealing with ‘clausal chaining’ in informal conversation, where long turns can exceed over a hundred words, and are comprised of multiple coordination. In summary it seems of questionable utility to measure the complexity of beginner learners’ production during communicative pair work tasks through the amount of subordination in AS-units; namely, because it is not a clear and prevalent characteristic of real-time spoken communication. It should also be noted that Berman and Slobin (1994) see subordination as a high-level skill that characterizes adult speech rather than young children, which in the case of the latter applies to beginner learners. It was for these reasons that I tried to think of different ways of conceptualizing language complexity. The next section will discuss how I conceptualized complexity in regards to my learners’ data.

Re-conceptualizing Complexity with Beginners and Open-ended Tasks

Bygate (2001) rather than working with subordination chose to look at words per unit as a measure of complexity. Bygate demonstrates with T-units (similar to AS-units) that complexity can be calculated by measuring the amount of words per T-unit. Besides being a broad fluency measure in relation to overall proficiency, Bygate argues that the number of words per unit reflects complexity in that the learner demonstrates the ability or at least attempts to combine lexical items around syntactic structures. Bygate (1999: 199) found that with argumentation and narratives tasks, learners, rather than embedding clauses resorted to increasing the number of clauses elements or words in their production, which meant that more complex clauses had a greater number of words. In addition word count as a measure of complexity is appealing because beginner L2 learners tend to process one word at a time when engaged in real time language production (see Foster 2001).

Departing from Mean Number of Words per AS-unit

With my data rather than Bygate's mean number of words per AS-unit, I instead categorized all AS-units in individual task performances into groups based on the number of words per unit. For example in each individual's performance all the five-word AS units were counted, and then all six-word AS-units were counted and so forth. By doing this I had a breakdown of the amount of various AS-unit lengths based on word count for each individual performance. This idea was derived from Skehan and Foster (2005) who measured error-free clauses while taking into account their word count. Their study provided a breakdown of accurate clauses based on their word count. The study established a 'cut-off' point where the learner past a certain threshold of clause length was unable to produce accurate clauses. Rather than accuracy I tried to do the same with complexity as AS-unit word count. The basic idea is that if a learner produces a majority of AS-units over an established benchmark based on word count, then the learner is recognized as 'stretching' their language production. The major question was what word count to use as the benchmark.

Establishing a Word length Benchmark for Complexity

Miller (1956), Chafe (1980), and Biber et al. (1999: 1067), contend that humans due to working memory limitations are only able to process six to seven words at a time during conversation. Skehan (2007) believes the limit to be even less (personal communication). Skehan's view is supported by Cowan (2001) who argues that it is four words. As Sinclair & Mauranen (2006: 35) note the precise number is unknown; however, the main point is that human processing of language is of limited capacity, and, therefore, speakers, especially during real time conversation, cannot be expected to produce overly elaborate turns. If this is

the case then it is clearly more limiting for L2 learners. After an initial analysis of the data, I concluded that the majority of AS-units in my data fall between a three and eight word range, before there is a significant drop at nine words and up. Taking into consideration the working memory issue just described above, and what I saw in the data, I assumed that a suitable benchmark might be around five to nine words. In order to verify this, my next step was to analyze AS-unit word counts. When I counted words per AS-unit, I did not consider grammatical correctness. For example *I living alone now* is missing the auxiliary verb, and I counted it as four words. If it had contained the auxiliary verb then I would have classified it as five words. As is common practice I did not include dysfluent language as part of an AS-unit. I counted any words that were part of the learners attempt to communicate, again minus dysfluencies. This meant that I counted *yes* and *no* as part of an AS-unit's word total when either was present. As is also common practice I counted contractions as two words, and, I counted Japanese words that were proper nouns like a city or place. Back channelling or simple repeating of a partner's words was not counted as part of a learner's production. In the excerpts and transcripts AS-units are demarcated by slashes: S1:/(It) it don't become round/(IAS 5). The brackets contain the number of AS units present in S1's turn, and then how many words the AS-unit contains.

One of the problems of analyzing units to look for different features is that learners of course have different interlanguages, as language development is a nonlinear process (Larsen-Freeman and Cameron 2008). Moreover, what I was trying to do was distinguish qualitative differences between unit word counts that differ from only one to four words. Taking this into consideration what I attempted to do was tease out very general, and tentative characteristics for different AS-unit word counts in order to provide myself, if possible, with some threshold where units from a certain word count become difficult to produce for all four learners in my data, and also demonstrate learner risking taking and language complexity. To do this I analyzed all AS-units in their respective word length groups. For example I reviewed all four-word AS-units and then all five-word units and so forth. I did not include question forms in my analysis, as I was only interested in looking at production oriented towards answering. When I analyzed the word count of units, I looked at general features, for example how many AS-units are error-free, and how many are full clauses as opposed to sub-clausal units. In addition I looked for the presence of subordinators and coordinators, and whether subordination or coordination is present within the unit, or whether the subordinators and coordinators are serving as adjuncts between units. I also looked for when a wider breadth of subordinators are present, like for example conditionals and relative clauses. Finally, I checked for other features like negation, the

presence or absence of adjectives, and the breadth of verb tenses at different word counts. Expectedly, learners did not produce an equal amount of for example four-word units, so I was careful not to over or under represent a learner's production and tried to weigh all learners' production equally.

General Characteristics of Different AS-unit Lengths

In the data with all the learners one to three-word AS-units are mainly sub-clausal units. These units are elliptical, minimal, and many are short exclamatory comments, or short questions and answers. However, a small number of the three-word units are basic clauses like *He is kind*. With four-word AS-units in the data some are sub-clausal units, but most are basic clauses with some having adjectives and adverbs. Most basic clauses at this word count in the data are error-free, and I found no negation present in any four-word AS-unit. The examples below are typical of what a majority of the four-word AS-units look like for all four learners. They are declarative statements in present tense and simple past tense.

Four-word AS-units

1. S1: /*(mm 3.0) Its name is Maru* /
2. S1: /*Rabbits hate rain water*/
3. S2: /*And second is money*/
4. S3: /*(My) my major is Denmark*/
5. S3: /*(That that is) maybe that is interesting*/
6. S4: /*I was very surprised*/
7. S4: /*I'm from Kyoto*/

With four-word units I concluded that the learners have no trouble with them, and these units are mostly simple clauses in the simple and past tense.

With five-word units and larger, I found a larger variety of simple verb phrases. Negation is also present from five-words and on.

Five-word AS-units

1. S1: /*(It) it don't become round*/
2. S1: /*And (1.60 mm) (I heard 3.30 mm it) ah I heard (4.50) (his his) his voice (laughs)*/
3. S2: /*She said it is fun (its very fun)*/
4. S2: /*Ah I think (12.00) over (over fou) four months*/

5. S3: /I want a big party/
6. S3: /But most important is character (au kado ka maybe)/
7. S4: /Recently (1.19) such crimes are increasing/
8. S4: /When she went to bowling/

Six-Word AS-units

1. S1: /(When) when it sleep it become (rou) round/
2. S1: /And (it likes to) it likes (1.57) to (.84) bite anything/
3. S1: /So it is (mm 1.53 old) old rabbit now/
4. S4: /(I don't know) I don't know any crimes/
5. S4: /Ah my neighbor was stolen her bike/
6. S4: /No no she asked but she said/
7. S3: /That is good point I think/
8. S3: /(I don't) I don't like who lie (who 2.26 lie)/
9. S2: /And second is (3.56 ah) he's tall/
- 10.S2: /It's very difficult for me/

A number of units at the five and six-word count are comprised of a *personal pronoun + verb phrase*. For beginner learners producing the appropriate pronoun and simple verb phrase may be easy; however, the real difficulty for them as Willis (2003: 71) notes is choosing the patterns that follow different verbs. The fact that a larger variety of simple verb phrases are more abundant suggests that learners would have to start tackling more difficult verb complementation from five-word units and longer. Example eight in the six-word group appears to be evidence of this. What is noticeable at the six-word AS-unit length is the presence of more units with errors and dysfluencies. Five-word units have dysfluencies, but in my data most are error free, as are four-word units. I concluded that it is from the six-word unit length that errors become more prevalent. As Skehan and Foster (2005) note this is to be expected when learners attempt to produce longer units that stretch their processing capabilities.

At the six-word length as the examples above show there are units with subordination and coordination occurring within them. In case of five-word units, I found no examples of subordination or coordination occurring within them. At both the five and six-word unit length, subordinators and coordinators are common at the beginning of the unit, for which their roles appear to be only as discourse markers. However this is misleading as in some cases as Carter and McCarthy (2006: 262) note coordinators besides acting as 'sentence

beginners' can also provide cohesive links between clauses, or in the case of this study AS-units. Many of the coordinators and subordinators acting as sentence starters are actually learners trying to structure a piece of production that transcends the boundaries of individual AS-units. In this sense they are more than discourse makers, but relate to a larger discourse grammar in the data.

Seven, eight, and nine word units are essentially the same as six-word units in that I found subordination and coordination occurring within units. (see examples below).

Seven-word AS-units

1. S1: /Mm because (it) it becomes round when it sleeps/
2. S4: /(Not um I) I through Takarazuka and went to Mukogawa/
3. S3: /But I want to honeymoon to abroad/
4. S2: /And (2.91 tsutashimi ga aru friendly) a small marriage (will be) will be friendly/

Eight and Nine-word units

1. S1: /Cat is (1.99 not) not dislike but (1.98) dog is better/
2. S1: /(So but) So if we raised outside (mm 1.74) it dies (early) early (1.65)/
3. S2: /Because a big wedding (1.27) take a lot of money /
4. S2: /(I) I don't want to married (si silence) silence man/
5. S3: /He (he) don't have money (we we could (2.22) we could live) we couldn't live/
6. S3: /Different point is whether I love him or not /
7. S4: /So when the sunsets (uh) there is very dark/
8. S4: /Of course because a girl kidnapped (and killed) was killed/

Finally, as with six-word units the examples above show again that error and dysfluencies are very common. In summary with the data above you could choose a complexity benchmark between six to nine words.

Settling on an AS-unit Benchmark for Complexity

After reviewing all units within the three to nine-word range, I decided on a seven-word AS-unit length as the benchmark for complexity in my data. A six-word length would also have been suitable, since every feature I identified in seven-word units is also present in six-word units. The choice of a seven-word benchmark makes the benchmark standard slightly higher. Table 1 presents the total number of six and seven-word units produced by each of the four learners that comprised my dataset.

Table 1. Total Numbers of Six- and Seven-word AS-units by Each Learner

	Kazue S1	Miho S2	Naoko S3	Rie S4
Six-word	11	8	10	12
Seven-word	11	10	10	11

For each learner, the sum of the six and seven-word totals represents just over 20% of the AS-units for their individual performances. Because of this fairly high percentage I decided to separate these two categories, despite their similarity. I didn't feel this was problematic for the benchmark concept because, as noted above, there appears to be a clear difference in the data between five-word and six-word units. Looking at the examples above, especially of eight and nine-word units, it seems clear that they indicate language complexity and risk-taking, and moreover stretch all four learners' interlanguage. In fact a nine-word length might have served as a benchmark, as the number of AS-units of nine words and above that learners produce drops substantially. However, I concluded that for these four learners nine words was too high.

Using the Benchmark

Below is an example of how you can use words per AS-unit and the benchmark to analyze learner performances. In the data below the learner was engaging in immediate task repetition of the same communicative task with different partners on the topic of crime. The first column is the number of words per AS-unit and the other three columns represent each of the three performance. For example we can see that the learner in her first performance produced zero AS-units of one word, one in the second performance, and five in the third performance.

Table 2. S3's AS-unit Breakdown

AS-unit	First Performance	Second Performance	Third Performance
1	0	1	5
2	2	0	6
3	3	3	4
4	1	4	3
5	1	2	2
6	4	1	5
7	3	3	4
8	3	1	0
9	2	2	1
10	2	4	3
11	1	1	3
12	1	0	0
13	0	0	0
14	1	0	1
15	1	0	0
16	3	0	0
17	2	1	0
18	0	0	0
19	0	0	0
20+	0	2 (24, 31)	0

Applying the benchmark concept to the data above for the learner S3 we can identify a trend in her performance of doing the same task three times.

Table 3. S3's AS-unit Benchmark

AS-units under 7 words	11/30 37%	11/25 44%	25/37 68%
AS-units of 7 words or more	19/30 63%	14/25 56%	12/37 32%

In the first performance 63% of AS-units are seven words and over. In addition 11 of these 19 units are 10 words and over. S3 produces three 16-word units and two 17-word units in her first performance. In contrast, in the third performance her long unit production drops considerably. While she produces 37 AS-units, 25 of these are under seven words; moreover, she produces only one unit of 12 words and over in the third performance. Fifteen of her units are three words or less. The measurement tool provides a clear picture of S3's production from the first to the third performance becoming more elliptical and minimal. In other words her task performance gets worse.

Conclusion

The purpose of this paper was to demonstrate the idea of a benchmark based on words per AS-unit as an acceptable option for measuring beginner learner's language complexity during real-time communicative pair work task. The number of words per AS-unit reflects the

ability of the learner to attach language around syntactical structure or follow an add on strategy by ‘tacking on’ language in a ‘bit by bit’ process. In regards to counting the number of words per AS-unit as a measure of complexity the credibility of the measure depends upon demonstrating that AS-units from a certain length start to contain features of language complexity that are not present in shorter units. With my data I suggested that from the six-word AS-unit and above subordinators, coordinators, and clausal embedding are present, and, moreover most units are comprised of complete clauses rather than being sub-clausal. I also demonstrated that at this length and above a larger breadth of tenses are present, and, furthermore, learner errors and dysfluency increase as learners processing capacity is stretched. I believe that AS-unit length combined with a benchmark can be an effective measure for identifying general changes in the learners’ production. I think its effectiveness is partially attributable to the fact that other than back channel language, I included almost all of the learners’ language to track the quality of their task performances. Even the small units provided valuable information about movement towards or away from minimal and elliptical language production. In conclusion word count and an accompanying complexity benchmark have sufficient descriptive potential for researchers who are conducting small qualitative studies.

References

- Berman, R. and Slobin, D. 1994. *Relating Events in Narrative: Crosslinguistic Developmental Study*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finnegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Brown, G. 1989. Making sense: The interaction of linguistic expression and contextual information. *Applied Linguistics* 10: 97-108.
- Bygate, M. 1999. Quality of language and purpose of task: Patterns of learners’ language on two oral communication tasks. *Language Teaching Research* 3: 185-214.
- Bygate, M. 2001. Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan and M. Swain (eds), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*. Harlow: Longman.
- Carter, R.A. and McCarthy, M. J. 2006. *Cambridge Grammar of English: Comprehensive Guide to Spoken and Written English Grammar and Usage*. Cambridge: Cambridge University Press.
- Chafe, W. 1980. The deployment of consciousness in production of a narrative. In W. Chafe (ed), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.
- Cowan, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage

- capacity. *Behavioral and Brain Science* 24: 87-185.
- Ellis, R. and Barkhuizen, G. 2005. *Analysing Learner Language*. Oxford: Oxford University Press.
- Foster, P., Tonkyn, A. and Wigglesworth, G. 2000. Measuring spoken language: A unit for all reasons. *Applied Linguistics* 21: 354-75.
- Foster, P. 2001. Rules and Routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan and M. Swain (eds), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*. Harlow: Longman.
- Larsen-Freeman, D. and Cameron, L. 2008. *Complex Systems and Applied Linguistics*. Oxford: Oxford University Press.
- McCarthy, M. J. 1998. *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.
- Miller, G. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81-97.
- Schleppegrell, M. 1992. Subordination and linguistic complexity. *Discourse Processes* 15: 117-31.
- Sinclair, J. and Mauranen, A. 2006. *Linear Unit Grammar*. Amsterdam: John Benjamins.
- Skehan, P. 2001. Tasks and language performance assessment. In M. Bygate, P. Skehan and M. Swain (eds), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*. Harlow: Longman.
- Skehan, P. and Foster, P. 2005. Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (ed), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.
- Willis, D. 2003. *Rules, Patterns and Words: Grammar and Lexis in English Language Teaching*. Cambridge: Cambridge University Press.