

Linear unit discourse analysis: The case of peer group interaction in the HKDSE public examination

Cameron Smart

Hong Kong Examinations and Assessment Authority

This article represents a further exploration of the viability of the model of Linear Unit Discourse Analysis (LUDA) proposed by Smart (2016) through the analysis of the spoken discourse of candidates in the peer group interaction component of the Hong Kong Diploma of Secondary Education (HKDSE) speaking public examination. LUDA is a linear description of analysis following the tradition of John Sinclair and his associates (e.g., Brazil, 1995; Sinclair, 1992, 1993/2004; Sinclair & Mauranen, 2006), which purports to describe the linguistic phenomena that signal the links between the sequence of incoming language as the listener experiences it. In this article, the HKDSE candidates' discourse is compared with the discourse of a university seminar discussion from the MICASE corpus using LUDA. It is found that HKDSE candidates' discourse tends to display little in terms of organizational-interactive linear units prevalent in the MICASE data. Neither does the HKDSE candidates' discourse contain much in the way of suspensive linear units between turns, indicating that there are relatively fewer cases of contradictions, disagreements and the type of convergent interaction that one would expect in such discourse. Instead, HKDSE candidates' turns tend to be lengthier and display more of the characteristics of written speeches.

Key words: linear unit discourse analysis; speaking examination; peer group interaction

Introduction: Linear approaches to the description of language

A linear description of language, as envisaged by John Sinclair and his associates (Brazil, 1995; Mauranen, 2012, 2016; Sinclair & Mauranen, 2006), is characterised fundamentally as one which describes language "as far as possible in a linear succession of units" (Sinclair & Mauranen, 2006, p. xix). Its aim is to describe how "[p]eople experience language as a linear phenomenon, that is, arranged along one dimension" (Sinclair & Mauranen, 2006, p. 5). This dimension is time in the case of spoken discourse and space in the case of written discourse. The result of such a dynamic perspective is a description of language which is inevitably synoptic (Mauranen, 2009b), i.e. one that describes the real-time unfolding of the discourse as the listener/reader experiences it, contrasting with a static description, typically seen, for instance, in conventional grammars, in which the text is examined as a completed object which "already exists in its entirety" (Brazil, 1995, p. 37).

In a linear approach, then, the researcher focuses on the "text of the moment" (Sinclair, 1993/2004, p. 82), i.e. the part of the text which s/he is engaged with at that moment and is asked to relate the text of the moment both to what has immediately preceded it in the text and to consider whether it leads the listener/reader to expect something in the immediate upcoming text. A linear description describes the sequence

of parts of the text and the constraints in place on these sequences on the syntagmatic axis as opposed to the variety of choices available on the paradigmatic axis.

The most well-known of these linear descriptions of language is Linear Unit Grammar (LUG) (Sinclair & Mauranen, 2006). LUG represents a radical departure from conventional grammatical descriptions in several different ways. The most relevant of these to this article is the fact that LUG employs an innovative system of analysis whereby elements are categorized as being message-oriented (M), organizational-textual (OT) or organizational-interactional (OI). M elements are then further subdivided into various categories (M-, +M, MS etc.) depending on whether they are considered to be complete, incomplete, supplementary and so on. LUG has thus far been employed in only a limited number of studies, in the study of spoken discourse used by second language users of English (Carey, 2013; Mauranen, 2009a) and in the study of online written discourse (Smart, 2016).

Linear Unit Discourse Analysis (LUDA) is a term coined by Smart (2016) to describe a system of analysis, similarly based on linear principles but designed to describe the relationships *between* Linear Units as opposed to the relationship between elements *within* Linear Units as is the case in LUG. It hence applies the same principles and general categorization proposed by Sinclair and Mauranen in LUG to the analysis of discourse and consequently employs the same general categorization (M, +M, OI etc.) seen in LUG. This is achieved through the incorporation of insights gleaned from Sinclair and his associates' earlier work in discourse analysis (e.g., Francis & Hunston, 1992; Sinclair, 1992, 1993/2004; Sinclair & Coulthard, 1975).

LUDA is fundamentally a linguistic description (Coulthard & Brazil, 1992). It is linguistic in that it is, based on a categorization of linguistic phenomena. As Sinclair (1992) argues, to reflect the experience of the language user, a linguistic system of analysis should reduce this experience to a small definable number of linguistic phenomena. In LUDA, this is achieved with reference to a small number of "syntagmatic mechanisms" (Smart, 2016) i.e. features of the incoming text that allow the reader/listener to create the structure of the discourse, such as prospection and encapsulation. These are used as a means of delimiting ten defined sub-categories of linear units (see Table 1). According to Sinclair and Coulthard (1975), such a categorization in a linguistic description should be capable of dealing with all of the discourse under examination. This contrasts with the more widely-known research in the field of discourse analysis that employs Conversation Analysis (CA) (see, for example, Jin & Zhang, 2015, for a review of the use of CA in recent research in oral assessment). CA has, nonetheless, met with criticism over the years for having its roots in the non-linguistic field of sociology and for employing a seemingly unlimited number of non-linguistic terms to describe specifically chosen points of interest (Coulthard & Brazil, 1992).

As of now, the only detailed study employing LUDA has been that of Smart (2016), which has as its focus the polylogic written discourse of online message boards. This article, therefore, represents a further stage in the validation of the LUDA model by exploring the insights that LUDA can bring to spoken discourse produced by learners of English in the context of the Hong Kong Diploma of Secondary Education (HKDSE) Paper 4 Speaking examination.

System of analysis: The model of Linear Unit Discourse Analysis (LUDA)

The unit of analysis, according to Smart (2016), in LUDA is the linear unit. It is defined as a meaningful segment of discourse constituting at least one core element plus

associated qualified elements (see Smart, 2016 for a full description of how the linear unit is delimited). There are two main categories of linear units: message-oriented (M) linear units and organizational-interactional (OI) linear units. These are divided between core and qualified-status linear units (see Table 1); qualified-status linear units being those which contain one or more syntagmatic mechanism.

The four syntagmatic mechanisms referred to in LUDA (Smart, 2016) are:

1. **Prospection:** The quality in the phrasing of the text of the moment whereby it “leads the addressee to expect something specific in the next part of the text” (Sinclair, 1993/2004, p. 88), whether it be some sort of elaboration or, in a dialogic discourse, an eliciting move in an elicit exchange (Francis & Hunston, 1992; Sinclair & Coulthard, 1975).
2. **Completion:** This signals the point in the text where what has been prospected previously is fulfilled. In discourse, this may be the completion of a list or a response to an eliciting move in a dialogue.
3. **Encapsulation:** A retrospective mechanism in which the previous text is referred to, encapsulated and incorporated into the present linear unit (Francis, 1994; Sinclair, 1992, 1993/2004). Explicit encapsulation (Sinclair, 1993/2004) is evident in anaphoric signalling noun phrases such as *that information* and *your question*; and in pro-form demonstratives e.g. *this, that*. According to Sinclair (1992, 1993/2004), encapsulation is also present in an array of more implicit devices from logical connectives and discourse markers to monosyllabic acknowledgements in a dialogue, e.g. *yeah, mhm*.
4. **Overlay:** Another retrospective mechanism where “the text of the moment contains elements that act as a paraphrase or an approximate antonym of the whole or part of the previous linear unit” (Smart, 2016, p. 56).

Qualified linear units are further sub-divided into compliant linear units, i.e. those that comply with the linear expectations of the discourse and suspensive linear units, i.e. those which do not¹ (see Table 1 for details). In addition, these linear units can be further divided in terms of whether they demonstrate relations within the same turn or with a previous or upcoming turn. These will be referred to as having *intra-turn orientation* and *inter-turn orientation* respectively.

To illustrate this categorization we can consider the properties of an M– linear unit (see L1 of Example 1). Firstly, it is message-oriented and therefore contributes to the incrementation of the shared knowledge between participants (Brazil, 1995; Sinclair & Mauranen, 2006). Secondly, there is no overt linguistic link evident in the linear unit to the previous linear unit. Thirdly, it is incomplete. This can mean two things depending whether the linear unit has intra-turn orientation or inter-turn orientation. If an M– linear unit has an inter-turn orientation (as is the case in Example 1), it means that an expectation is created that another participant should complete the uncompleted linear unit. This has previously been described in terms of representing the first move in an elicitation-response exchange (Sinclair & Coulthard, 1975).

Table 1. Subcategories of linear units

Type	Symbol	Description	Syntagmatic mechanism(s) present
Core status M linear unit	M	Complete linear unit of meaning which displays no structural link with previous or upcoming linear unit	~
Qualified status compliant M linear unit	M–	Incomplete linear unit of meaning which displays no structural link with previous linear unit	Prospection
	+M	Linear unit of meaning which completes previous incomplete linear unit of meaning	Completion
	+M–	Incomplete linear unit of meaning which partially completes previous incomplete linear unit of meaning	Prospection & completion
	MS	Supplement to linear unit of meaning	Encapsulation or overlay
	MS–	Incomplete supplement to linear unit of meaning	Encapsulation or overlay and prospection
Qualified status suspensive M linear unit	M	A complete linear unit of meaning which does not comply with linear expectations	Encapsulation or overlay
	M–	An incomplete linear unit of meaning which does not comply with linear expectations	Encapsulation or overlay and prospection
Qualified status compliant OI linear unit	OIS	Interactional supplement to linear unit of meaning	Encapsulation or overlay
Qualified status suspensive OI linear unit	OI	A linear unit of interaction which does not comply with linear expectations	Encapsulation or overlay

Example 1: M– / +M sequence with inter-turn orientation²

P ³	T	L	Linear unit	Type
A	1	1	what is your opinion?	M–
B	2	2	I think the scheme eh is benefit to save the energy	+M

In an M– linear unit with an intra-turn orientation, the listener perceives some sort of commitment on the part of the speaker to provide a completion in the next linear unit.

Hence in Example 2, for instance, the participant commits to providing an elaboration of the first linear unit in L1. This is made explicit by the phrase *not only*. This prospection is duly completed in L2.

Example 2: M– / +M sequence with intra-turn orientation

P	T	L	Linear unit	Type
A	1	1	Eh for me when I writing diaries you can draw some pictures not only about the entrance ticket	M–
		2	we can also feel my teenage life is more meaningful	+M

In order to illustrate the difference between compliant and suspensive linear units, Example 1 can be contrasted with Example 3, which illustrates an occasion when the second participant does not comply with the linear expectations, i.e. does not complete the prospection.

Example 3: M– / |M sequence with inter-turn orientation⁴

P	T	L	Linear unit	Type
A	1	1	Would you like to see the Chinese white dolphins in the wild?	M–
B	2	2	Sorry can you change the other question.	M

As can be seen, the candidate (Participant B) does not comply by providing a response to the elicitation introduced by the examiner (Participant A). Instead, the expectations are suspended and the question remains unanswered. This is represented by the |M linear unit seen in L2.

Method

Data Background

The HKDSE Paper 4 Speaking paper, which has been administered since 2012, is the latest in a long line of English Language oral examinations which have been taken by secondary school students in Hong Kong since the 1950s. The paper comprises two components: a public examination component and a School-based Assessment (SBA) component. Both of these include a peer group interaction component. The peer group interaction component of the HKDSE SBA has attracted attention in recent years, with some varied conclusions regarding the quality of interaction produced by candidates.

Employing CA, Gan, Davison, and Hamp-Lyons (2009) and Gan (2010) found that peer group interaction provided the students with an opportunity to display a wide range of communication strategies. In contrast, (Luk, 2010, p. 25), also employing CA, found that the discourse produced in SBA was characterised as being “ritualized, contrived, and colluded...suggest[ing] a strong desire on the part of the students to maintain the impression of being effective interlocutors for scoring purposes rather than authentic communication”. Similarly, (Lam, 2015) found that although the discourse in the group interaction of SBA at first sight seems to display the features of natural interaction, it does in fact reflect significant rehearsal for the performance on the part of the students and as such the amount of natural spontaneous interaction present in the discourse between students was found to be limited.

Rather than focusing on SBA, this article represents an investigation of candidates’ interaction in the peer group interaction component of the HKDSE in the public examination setting. The data is taken from transcripts of candidates’ performance, which have been routinely recorded by the Hong Kong Examinations and Assessment Authority since 2011, though not available to the general public. This article is the first to be published employing an extended set of data comprising candidates’ actual speaking performance in a public examination in Hong Kong.

Tagging Method

This study, then, is based on the comparison of two sets of data. The first (hereafter the DSE discourse) constitutes a transcript of different candidates’ performance in the peer group interaction component of the public examination component of the HKDSE Speaking paper, in which typically four candidates are given eight minutes to discuss a given topic having been given ten minutes preparation time. The DSE discourse comprises 8,457 words and represents the contributions of twenty-five different HKDSE candidates in twenty-five separate peer group interactions. The twenty-five candidates comprise five candidates from each of the five levels⁵ thus covering the spectrum of proficiency in the HKDSE.

The second set of data is taken from the Michigan Corpus of Academic Spoken English (Simpson, Briggs, Ovens, & Swales, 2002) (hereafter the MICASE discourse). It comprises a group of four undergraduates (two native speakers of English and two non-natives) in a first-year philosophy seminar at the University of Michigan, USA. It comprises 2,024 words. It was selected as being a highly interactive piece of discourse which reflects the type of real world spoken discourse that the HKDSE peer group interaction simulates. It was hoped that the comparison between the DSE discourse and the MICASE discourse may highlight certain linear sequences that were relatively prevalent or absent in the DSE discourse when compared to a similar real world discussion.

These two sets of data were tagged by the author following the system of analysis set out in Table 1 using the *UAM CorpusTool* (O'Donnell, 2009). An intra-rater reliability procedure was carried out whereby the author carried out the tagging of the data four weeks after the initial tagging. This produced a Kappa co-efficient of 0.87. Following the standard interpretation of the Kappa co-efficient (Landis & Koch, 1977), a figure of 0.81 or above signifies almost perfect agreement between raters. It was therefore concluded that the tagging was stable enough to proceed.

The frequencies of the occurrences of the LUDA categories and the subcategories in the DSE discourse and the MICASE discourse were then compared. The Log Likelihood Ratio (G^2) (Dunning, 1993) was used in order to calculate whether the

difference between the actual observed frequency of a category or subcategory of the linear in the DSE discourse was statistically significant when compared to what the expected frequency would be if they were distributed in the same way as in the MICASE discourse.

If no significant difference was found between the observed and expected frequencies then it is concluded that there was no difference between the two sets of data in that particular aspect. If a significant difference is found, it is judged that the linear unit subcategory is over-represented or under-represented in the DSE discourse when compared to the MICASE discourse⁶.

The p value in such calculations was set at $p < 0.001$ for an over/under-representation of high significance meaning a critical value = 10.83 (indicated in the tables in this article by **). Additionally, the p value was set at $p < 0.01$ for an over/under-representation of low significance meaning a critical value = 6.63 (indicated in the tables in this article by *).

Results and discussion

Table 2 shows the significance of difference in the distribution of M and OI linear units across the DSE and the MICASE discourse. In the second and third columns it shows the observed frequencies in the two sets of data and in the fourth and fifth columns, it shows the expected frequencies of the categories in the two sets of data as calculated, using the log-likelihood ratio. The final column shows the log-likelihood ratio value (G^2). As can be seen in the table, there is no over-representation or under-representation of M linear units evident in the DSE discourse ($G^2 = 2.63$).

Table 2. DSE vs. MICASE linear unit categories

	Observed frequencies		Expected frequencies		Log likelihood (G^2)
	DSE	MICASE	DSE	MICASE	
M linear units	449	246	428.32	266.68	2.63
OI linear units	20	46	40.68	25.32	-26.52**

However, there is a significant under-representation of OI linear units in the DSE discourse ($G^2 = -26.52$), meaning that OI linear units are much more prevalent in the MICASE discourse. An example of such an OI linear unit in the MICASE discourse can be seen in Line 2 of Example 4:

Example 4: M / OIS sequence with inter-turn orientation

P	T	L	Linear unit	Type
A	1	1	I thought this was the best, the best reading	M
B	2	2	yeah it is yeah yeah	OIS

According to Sinclair (1992), even a minimal acknowledgement as can be seen in L2 is an instance of encapsulation of the previous linear unit, albeit an implicit encapsulation, in that it encapsulates the ongoing success of the discourse up to this point. This type of M / OIS sequence between turns plays only a minor role in the DSE discourse.

Despite there being no over- or underrepresentation in general in M linear units, there is significant underrepresentation in certain subcategories (see Table 3) in the DSE discourse. Most strikingly, there is an underrepresentation of suspensive |M linear units ($G^2 = -13.51$) and in retro-inter turn M linear units ($G^2 = -34.53$) in the DSE discourse (see Table 4). Taking these two results together, this indicates that there are significantly fewer message-oriented linear units which in some way flout the discourse expectations between turns in the DSE discourse when compared to what would be expected if they occurred at the same rate as in the MICASE discourse.

Table 3. DSE vs. MICASE M linear units

	Observed frequencies		Expected frequencies		Log likelihood (G^2)
	DSE	MICASE	DSE	MICASE	
Core status M linear unit	31	14	27.73	17.27	1.03
M- linear unit	28	6	20.95	13.05	6.91*
+M linear unit	77	34	68.41	42.59	2.90
+M- linear unit	25	6	19.11	11.89	5.23
MS linear unit	162	88	154.07	95.93	1.07
MS- linear unit	69	30	61.01	37.99	2.81
M linear unit	9	22	19.11	11.89	-13.51**
M- linear unit	32	28	36.98	23.02	-1.71

Table 4. DSE vs. MICASE M linear unit orientation

	Observed frequencies		Expected frequencies		Log likelihood (G^2)
	DSE	MICASE	DSE	MICASE	
M pro-intra turn	110	48	97.37	60.63	4.40
M pro-inter turn	44	22	40.68	25.32	0.72
M retro-intra turn	296	94	240.35	149.65	35.87**
M retro-inter turn	78	114	118.33	73.67	-34.53**

An example of a suspensive |M linear unit with inter-turn orientation taken from the MICASE discourse can be seen in L3 of Example 5, where Participant B disagrees with

what has been said by Participant A. This is signalled by the organizational-interactional element *yeah but*. As noted above, this type of implicit encapsulation acknowledges the participant has understood the discourse so far. In this case, however, it also indicates that s/he has evaluated the previous discourse and disagrees with it. This flouting of the expectation of agreement means that it is characterised as being suspensive. This is then reiterated in the continuation of Line 3 by an overlay (Sinclair, 1993/2004; Smart, 2016), seen in the slight adaptation of *if someone asked you* (L1) to *if someone asks you* (L3). A further example of a suspensive inter-turn |M element can be seen in L5 in the same example. In this case, in L4, Participant C completes Participant B's incomplete statement in L3. However, Participant B is unhappy with this completion and in L5 corrects Participant C, thus employing a suspensive linear unit. To do this, Participant B uses an overlay through the use of a parallel structure within which s/he replaces *experiencing* with *thinking*. Realizing the sensitivity of correcting Participant C, Participant B ends the linear unit with an echoing of *yeah* as seen in Participant C's previous linear unit.

Example 5: |M linear unit with inter-turn orientation in co-constructed discourse

P	T	L	Linear unit	Type
A	1	1	yeah but then if someone asked you while you were sitting there, where are you? <LAUGH>	M–
		2	you're not gonna be like	+M–
B	2	3	yeah but, if someo- if someone asks you that's a little different that what you're	M–
C	3	4	yeah experiencing yeah	+M
B	4	5	thinking during it yeah.	M

As can easily be appreciated, such sequences, containing inter-turn suspensive linear units, lead to a co-constructed meaning from the shared interaction between participants. These are very few and far between in the DSE discourse. Indeed, any inter-turn suspensive linear units are extremely rare in the DSE discourse and when they do occur they tend to be occasions when a candidate decides to try to clarify what another weaker candidate has said previously. This can be seen in Example 6.

Example 6: |M linear unit with inter-turn orientation for clarification

P	T	L	Linear unit	Type
A	1	1	Less than less: (5)	M
B	2	2	Maybe you want to say about the health problem	M–
		3	since mooncakes have high calories so we sent only one mooncake is best.(3)	+M

Example 7: |M linear unit with inter-turn orientation for contradiction

P	T	L	Linear unit	Type
A	1	1	because Hong Kong are too small	MS
B	2	2	But -but there is space enough right? I think. (3)	M-
		3	Anyway, another reason I want to say is that is uhm I want to say I think that it is the demand is uh equal price opportunities for Hong Kong musicians	M

There is only one occasion in the entire DSE discourse data (in L2 in Example 7) when a participant actually spontaneously disagrees with what another participant has said through the use of a suspensive |M linear unit. This example is also significant because Participant A does not respond to Participant B's elicitation in L2. Presumably such a genuine response was so unexpected that Participant A does not know how to respond. This is then followed by an awkward three-second silence in L2, which Participant B eventually breaks by going into his pre-planned point in L3, thus suspending the prospected completion which was expected from Participant A. The opportunity for actual genuine interaction between the two candidates then passes.

In contrast, as seen in Table 4 above there is an overrepresentation of retro intra-turn M linear units in the DSE discourse ($G^2 = 35.87$), i.e. those which relate back to the previous linear units within the same turn. This is due mainly to the fact that DSE discourse has a large number of lengthier turns when compared to the MICASE discourse. An example of this type of lengthier turn in the DSE discourse can be seen in Example 8.

Example 8: A lengthier turn in DSE Discourse

P	T	L	Linear unit	Type
A	1	1	It may threaten the harmony of Hong Kong society.	M
B	2	2	Yes I agree with you.	MS-
		3	Most people think that they- the cosplay may waste time waste money	M-
		4	because the costumes it is many money to buy or to make them.	+M
		5	And then they think that the cosplayers may want to may just want to get attention from others.	MS
		6	Plus some people think that this their this is their style or their hobby	MS
		7	and maybe respect to them or just think they are (??) enough.	MS
		8	What are your views?	M-

In this example, a strategy employed by candidates already described by Lam (2015) is evident, whereby a candidate initiates the turn in L2 with a pre-planned statement of agreement, thus providing some sort of semblance of coherence with the previous turn (L1) through an MS– linear unit. However, instead of elaborating on his/her agreement, as is prospected, the candidate suspends this prospection and instead embarks on a separate pre-planned topic in L3. This new topic is elaborated on in L5–7, with a series of MS linear units, each beginning with a logical encapsulating device (Sinclair, 1993/2004): *and* (L5 and L7) and *plus* (L6). As Sinclair and Mauranen (2006) argue, discourse with a large number of this type of organizational-textual devices is typical of rather formal written discourse or prepared public speeches rather than what you might expect in an unplanned spontaneous spoken interaction among four participants.

Conclusion

This article represents further validation of LUDA as a viable model, which, when applied, produces insights about both the relative character of the discourse in terms of it being message or interactional in its orientation as well as how it is sequenced both within and between turns. In the case of the DSE discourse it has been shown that candidates display a well-developed ability to control the interaction through longer speech-like turns. These are heavily reliant on a succession of message-based linear units linked by textual-organizational elements and thus display more of the characteristics of written language than any significant presence of those organizational-interactional elements normally associated with spontaneous dialogue. In Mauranen's (2009a) terms, there seems to be some evidence to suggest that candidates in the peer group interaction of the public examination component of the HKDSE display the means of achieving divergent interaction, i.e. they are capable of taking hold of the discourse and moving it towards their own goals. In contrast, there is little evidence of candidates displaying an ability to engage in convergent interaction, i.e. working together to negotiate and co-construct meaning through the use of interactive convergent devices (such as those seen in the MICASE discourse in Examples 4 and 5 above).

It should be kept in mind that the amount of data employed in this study is small so any conclusions must remain tentative. Nevertheless the results from the linear analysis carried out for this article would seem to point towards the evolution of a distinct discourse-type in the peer group interaction component in the HKDSE with its own linear structure and interactional norms. The reasons why this may be the case are manifold, and can now be explored further based on the evidence provided in this article. Following from this study, LUDA can now be employed for further research comparing the discourse in the HKDSE public examination context in Hong Kong with the peer group interaction of SBA. LUDA may also be employed in exploring whether these results are consistent across the whole candidature or vary depending on the grade candidates achieve.

The above provides further validation of LUDA as a viable model of discourse analysis, which successfully meets Sinclair's (1992) challenge to reduce the vast complexity of human communication to a small number of simple linguistic phenomena. It is hoped that other researchers can now go on to apply the model to a variety of spoken discourse types, contexts and genres in order to further explore the insights that a linear description of discourse can bring.

Notes

1. In the coding of linear units, a dash (–) indicates an incomplete linear unit; a plus (+) indicates a completing linear unit; and a bar (|) indicates that the linear unit is suspensive.
2. Examples 1 & 2 are taken from the peer group interaction component on the HKDSE speaking public examination.
3. In example tables in this article, P=participant, L=Linear unit and T=Turn. A dotted line between linear units means it is a continuation of the same turn; double solid line mean a new turn.
4. Example 3 is taken from another part of the HKDSE speaking examination, where an examiner asks an individual candidate a question.
5. The five levels refer to the final grade awarded to the candidate for the paper. These are arrived at by the examiners based on a variety of criteria namely: pronunciation, communication strategies, vocabulary and language patterns and ideas and organization.
6. No evaluation is attached to the terms *over* and *under-representation* in this article. They simply refer to the statistical properties of being significantly more or less frequent.

About the author

Cameron Smart is a Manager (Assessment Development) at the Hong Kong Examinations and Assessment Authority. He has been working in the field of EFL teaching and testing for the last 29 years. He has a PhD in Applied English Linguistics from the University of Birmingham. His interests lie in the application of linear unit grammar and linear unit discourse analysis.

References

- Brazil, D. (1995). *A grammar of speech*. Oxford: Oxford University Press.
- Carey, R. (2013). On the other side: formulaic organizing chunks in spoken and written academic EFL. *Journal of English as a Lingua Franca*, 2(2), 207–228.
- Coulthard, M., & Brazil, D. (1992). Exchange structure. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 50–78). London: Routledge.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), 61–74.
- Francis, G. (1994). Labelling discourse: An aspect of nominal-group lexical cohesion. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 83–101). London: Routledge.
- Francis, G., & Hunston, S. (1992). Analysing everyday conversation. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 123–161). London: Routledge.
- Gan, Z. (2010). Interaction in group oral assessment: A case study of higher- and lower-scoring students. *Language Testing*, 27(4), 585–602.
- Gan, Z., Davison, C., & Hamp-Lyons, L. (2009). Topic negotiation in peer group oral assessment situations: A conversation analytic approach. *Applied Linguistics*, 30(3), 315–344.
- Jin, Y., & Zhang, L. (2015). The impact of test mode on the use of communication strategies in paired discussion. In G. Yu & Y. Jin (Eds.), *Assessing Chinese learners of English: Language constructs, consequences and conundrums* (pp. 61–84). Basingstoke: Palgrave Macmillan.
- Lam, D. M. K. (2015). Contriving authentic interaction: Task implementation and engagement in school-based speaking assessment in Hong Kong. In G. Yu & Y. Jin (Eds.), *Assessing Chinese learners of English: Language constructs, consequences and conundrums* (pp. 38–60). Basingstoke: Palgrave Macmillan.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174. doi: 10.2307/2529310
- Luk, J. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly*, 7, 25–53.
- Mauranen, A. (2009a). Chunking in ELF: Expressions for managing interaction. *Journal of Intercultural Pragmatics*, 6(2), 217–233.
- Mauranen, A. (2009b). Spoken rhetoric: How do natives and non-natives fare? In E. Suomela-Salmi & F. Dervin (Eds.), *Cross-linguistic and cross-cultural perspectives on academic discourse* (pp. 199–218). Amsterdam: John Benjamins.
- Mauranen, A. (2012). Linear Unit Grammar. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 3409–3417). Oxford: Wiley-Blackwell.

- Mauranen, A. (2016). Temporality in speech – Linear Unit Grammar. *English Text Construction*, 9(1), 77-88. doi: 10.1075/etc.9.1.05mau
- O'Donnell, M. (2009). The UAM CorpusTool: Software for corpus annotation and exploration. In C. Bretones Callejas, J. F. Fernández Sánchez, J. R. Ibáñez Ibáñez, E. García Sánchez, E. Cortés de los Ríos, S. Salaberri, S. Cruz, N. Perdú, & B. Cantizano (Eds.), *Applied linguistics now: Understanding language and mind / La lingüística aplicada hoy: Comprendiendo el lenguaje y la mente* (pp. 1433–1447). Almería: Universidad de Almería.
- Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). *The Michigan corpus of academic spoken English*. Ann Arbor, MI.: The Regents of the University of Michigan.
- Sinclair, J. M. (1992). Priorities in discourse analysis. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 79-88). London: Routledge.
- Sinclair, J. M. (1993/2004). Written discourse structure. In J. M. Sinclair & R. Carter (Eds.), *Trust the text: Language, corpus and discourse* (pp. 82–101). London: Routledge
- Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse*. Oxford: Oxford University Press.
- Sinclair, J. M., & Mauranen, A. (2006). *Linear unit grammar: Integrating speech and writing*. Amsterdam/Philadelphia: John Benjamins.
- Smart, C. (2016). *Discourse reflexivity in Linear Unit Grammar: The case of IMDb message boards*. Amsterdam/Philadelphia: John Benjamins.