The Effect of Task Design on Interactional Authenticity in TBLT:

A Mixed Methods Conversation and Register Analysis

Daniel Keller

Northern Arizona University

Abstract

The effect of task design features on interactional authenticity in L2 Task-based Language Teaching (TBLT) was investigated using a corpus of transcripts of ten different language learning tasks. Discriminant analyses were used to identify linguistic features which accurately predicted (>90%) which of the ten tasks each transcript belonged to. Prototypical transcripts for each task were then identified and analyzed using conversation analytic techniques to describe the authenticity of the interaction of each task. Task design features were then examined in light of the conversation analysis findings. Results suggest that participants' familiarity with the situation of the task and the content of the task, as well as the degree to which exchanging information is emphasized, affected the interactional authenticity of the conversations.

Keywords: Task-based Language Teaching, TBLT, interactional authenticity, authentic tasks, conversation analysis, register analysis.

The Effect of Task Design on Interactional Authenticity in TBLT:

A Mixed Methods Conversation and Register Analysis

Background

An important characteristic of tasks in Task-Based Language Teaching is authenticity. Early proponents of TBLT believed that tasks should represent real language use (e.g., Skehan, 1996). More recent thinking on the topic suggests that tasks should elicit interaction similar to interaction in non-pedagogical tasks even if the tasks themselves are not particularly realistic (Ellis, 2009). It has been argued, however, that task-based interaction necessarily lacks authenticity because students must meet pedagogical goals imposed by the classroom situation as well as the interactional goals imposed by the task workplans (Seedhouse, 2004).

Though this view is grounded in conversation analysis, it aligns with longstanding thinking on register variation. Biber and Conrad (2009) argue that linguistic variation is driven by variation in characteristics of the social situations in which language is used. In this view, the classroom situation and the situation created by the task compete to shape the language students use during task-based interaction. Thus, interaction becomes less authentic when it reflects the situational parameters of the classroom situation (Seedhouse, 2004) and more authentic when it reflects the situational parameters of the task situation (Ellis, 2009). For TBLT to successfully produce authentic interaction then, tasks should be designed in such a way that participants are able to respond to the situational pressures of the tasks, rather than the situational pressures of the classroom. To date, however, no research has investigated the impact of task-design features on interactional authenticity conceived in this way.

Research Questions

1. What linguistic and interactional features are associated with classroom interaction? How do these features vary across tasks with different situational pressures? Can this variation be tied to features of the task workplans?

2. What features of tasks prompt participants to produce more or less interactionally authentic language rather than the language of classroom interaction?

Method

As this is an exploratory study with no body of existing literature to draw from, determining which task-design variables to manipulate experimentally and which linguistic features to measure as response variables was not possible. Therefore, rather than designing tasks and collecting new interaction data, the researcher chose to examine learner corpus data in a bottom-up fashion. This study used the Corpus of Collaborative Oral Tasks (CCOT) - a corpus of transcribed interactions between learners completing interactive oral assessment tasks (Crawford & McDonough, 2014). The corpus consists of more than 200,000 words of spoken data from 820 interactions (sound files and transcriptions) across 27 interactive oral assessment tasks. All tasks were completed in the NAU PIE between 2010 and 2014.

Using linear discriminant analysis (LDA; Izenman, 2103), lexicogrammatical features that reliably predicted the task that a transcript belonged to were identified. After several rounds of testing, the set of features was reduced from 150+ to 26. The reduced set of 26 features was able to classify texts with greater than 90% accuracy. Canonical discriminant analysis (CDA; Egbert & Biber, 2016) was then used to discover latent dimensions of variation in the frequency of these 26 features by task. The CDA identified nine dimensions of variation, but the first four

accounted for more than 80% of the total variance, so the remaining five dimensions were

excluded from the analysis.

The ten tasks with the most transcripts were selected for more detailed exploration.

Thirty-four transcripts from each of these ten tasks were randomly selected to produce a down-

sampled corpus of 340 texts divided evenly across ten tasks. Each text was then given a score for

each of the four dimensions of the CDA. From these text scores, average scores for each of the

four dimensions were then calculated for each of the ten tasks. These scores may be thought of as

coordinates for the tasks' centroids in four-dimensional space. Mahalanobis distances (De

Maesschalck, Jouan-Rimbaud, & Massart, 2000) from the tasks' centroids to the coordinates

(dimension scores) for each text were then calculated to identify which texts were closest to their

task centroids and thus prototypical for the task.

For each task, the four texts with the smallest distance values (and thus highest

prototypicality) were selected for close conversation analysis (CA) following principles outlined

in Heritage (2005), Markee, (2000), and Seedhouse (2005). Though the CDA identified

dimensions of variation in lexicogrammatical structures, the CA did not focus on those features

only, but rather took them as a jumping off point from which to examine those dimensions in

terms of variation in traditionally important CA constructs: turn taking, sequence organizing,

repair, and topic management, as well as task completion. These social actions were then linked

to situational parameters associated with task design variables and classroom interaction more

generally. Texts were read multiple times over the course of eight analyses – one for each pole of

the four dimensions. In many cases, the analysis of one pole of one dimension would support or

complicate an analysis of a previous pole of a previous dimension. As a result, the analyses

became both recursive and iterative as each pole stimulated reassessment of previous poles, which stimulated yet more reassessment.

Results

The results are organized as follows: A brief description of the typical classroom interaction is presented. Then each of the four dimensions is described. Finally, a description of task characteristics that align with variation along these dimensions is provided.

The Language of Task-based Interaction

All ten tasks followed a similar workplan. Students were given information or asked to brainstorm ideas or opinions on a topic. Then, students were instructed to communicate that information to their partner. Next students were asked to combine their information with their partners' and then make some sort of evaluative decision. For example, in one task, each student was given information about a location in a city. Then the students were asked to exchange the information they had and decide which location would be a more appropriate place to open a business.

Not surprisingly, this general structure was reflected in many of the transcripts. Students completed the task in roughly four phases: an opening phase, an information exchange phase, an operation phase, and a consensus phase. In the opening phase, students oriented to the task and established their roles. One student would typically take control of the interaction and initiate the next phase. In the information exchange phase, students exchanged the information or opinions either in single long turns, or in sequences of shorter turns. In the operation phase, students combined, rephrased, evaluated, or responded to the information exchanged in the preceding phase. Finally, in the consensus stage, students came to agreement on a single choice and closed the conversation.

Variation by Task

Dimension 1 – Other vs Self Agency.[1] The positive pole of this dimension includes features for describing the situations or actions of other people (pronouns, proper nouns) and commenting on those experiences (modals of necessity, stance nouns). The negative pole of this dimension includes an analogous set of features for performing the same functions in tasks that give agency to the task participants, rather than to a third party.

Dimension 2 – Specific vs General Reference. The features loading on the positive pole of this dimension were used to describe specific places, people, and things. This contrasts with tasks with features which load on the negative pole. These features allow students to discuss their topic in more general terms.

Dimension 3 – Informational vs Evaluative Interaction. Dimension 3 encompasses a split between exchanging information and evaluating information. Features loading onto the positive end of the pole were more informational and those loading onto the negative pole were more evaluative. As all tasks required both information exchange and evaluation of some kind, dimension 3 scores reflect the extent to which students had enough relevant background knowledge, opinions, and interest to focus on evaluation rather than information exchange.

Dimension 4 – Sparse vs Dense Information Flow. In general, the negative pole points toward informational, factual language, but aligns well with the 'literate discourse' pole of the Oral vs Literate Discourse dimension in Biber & Conrad's (2009) multidimensional analysis of university registers. Here, the dimension signals information density, with tasks at the negative pole having dense information structures and tasks at the high end having sparse.

Linguistic features loading onto the poles of each dimension appear in the Appendix.

---

[1] The author is grateful to William Crawford, Jesse Egbert and Soo-Jung Youn for help with interpreting these dimensions.

Task Features and Register Variation

Variation in the dimensions above can be linked directly to task features. These differences are summarized in Table 1.

Table 1

Design Features of Tasks Affecting Dimension Scores

Dimension	Task design features by dimension scores	
	High	Low
1. Other vs Self	Cooperating	Adversarial
2. Specific vs General	Familiar Situations	Unfamiliar Situations
3. Inform vs Evaluate	Unfamiliar Content	Familiar Content
4. Dense vs Sparse	Information Exchange is Central	Information Exchange is Peripheral

In terms of task design, dimension 1 was related to whether the task required students to adopt an adversarial or cooperating relationship. For example, were they asked to argue for and make a decision between one of two binary options or were they asked to consider a situation and propose a course of action together. Dimension 2 was related to how familiar the task situation was with more familiar situations evoking more specific language. Dimension 3 was related to how known the informational content of the task was. When students were asked to talk about topics they knew well, they produced more evaluative language. Finally, dimension 4 was related to how important the information exchange was to completing the task.

Relevance to PIE and Second Language Learning

The findings suggest that it is possible to design tasks that produce more interactionally authentic conversations in task-based interaction by manipulating their features. Tasks which scored low on dimension 2, and high on dimensions 3 and 4 tended to produce transcripts with long, non-interactive information exchange phases and minimal operation and consensus phases. These tasks encouraged students to focus on completing the assignment efficiently rather than

appropriately. Thus, they reflect the situational pressures of the classroom environment rather than the task. Often, there is little evidence in these transcripts that students understood one another as it was possible to complete tasks using knowledge of the workplan rather than through meaningful completion of the task.

Conversely, tasks which scored low on dimension 2 and high on dimensions 3 and 4 produced transcripts in which interaction was much more authentic. Information exchange proceeded across several turns with students offering backchannels, requesting clarification, and engaging with the information as it was introduced into the conversation. Similarly, the operation and consensus phases of these interactions were longer and more prominent in the conversation. These transcripts tended to include significant evidence that students understood one another and worked together to complete the task. Dimension 1 did not seem to impact the authenticity of students' interaction.

These findings suggest that PIE task-designers should consider the familiarity of the situation and content during task design, or that PIE teachers should spend time preparing students for tasks by allowing them to familiarize themselves with the task situation and content before being asked to do the task. On the other hand, the results also suggest that the longstanding division between open-ended and closed-ended tasks and between cooperating and competing tasks (e.g., Pica, Kanagy, & Falodun, 1993; encoded here in dimension 1) do not impact the interactional authenticity of a task.

The method used in this study is a promising way of investigating systematic differences in task design and language use in TBLT, but the current study suffers from a number of limitations. First, the ten tasks are quite similar in their designs. Features of the tasks that don't vary, or which don't vary much will not affect linguistic behavior and so remain hidden to the

CDA. Second, as this is an observational study, the findings must be considered tentative unless

replicated with a proper experimental design. Third, very little information is known about the

content of classes leading up to the tasks. These limitations should be addressed in future

research.

References

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge, UK: Cambridge
    University Press.

Crawford, W., & McDonough, K. (2014). *Corpus of Collaborative Oral Tasks*.

De Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L. (2000). The Mahalanobis distance.
    Chemometrics and Intelligent Laboratory Systems, *50*(1), 1-18.

Egbert, J., & Biber, D. (2016) Do all roads lead to Rome?: Modeling register variation with
    factor analysis and discriminant analysis. *Corpus Linguistics and Linguistic Theory*.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International
    Journal of Applied Linguistics*, *19*, 221-246.

Heritage, J. (2005). Conversation analysis and institutional talk. In K. L. Fitch & R. E. Sanders
    (Eds.), *Handbook of language and social interaction* (pp. 103-147). Mahwah, NJ: Lawrence
    Erlbaum.

Izenman, A. J. (2013). Linear discriminant analysis. In *Modern multivariate statistical
    techniques* (pp. 237-280). Springer New York.

Markee, N. (2000). *Conversation analysis*. Mahwah, NJ: Lawrence Erblaum Associates Inc.

Pica, T., Kanagy, R., Falodun, J. (1993). Choosing and using communicative tasks for second
    language instruction. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning:
    Integrating theory & practice*. Multilingual Matters.

Seedhouse, P. (2004). Conversation analysis methodology. *Language Learning*, *54*(S1), 1-54.

Seedhouse, P. (2005). Conversation analysis and language learning. *Language Teaching*, *38*,
    165-187.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied
    Linguistics*, *17*, 38-62.

Appendix

Dimensions and the Features that Load on them.

Dimension 1		Dimension 2		Dimension 3		Dimension 4	
Predictor	Score	Predictor	Score	Predictor	Score	Predictor	Score
3rd person pronouns	**0.879**	**place nouns**	**0.638**	**concrete nouns**	**0.427**	2nd person pronouns	0.273
communication Verbs	**0.444**	**nominalizations**	**0.612**	**passives**	**0.339**	pronoun 'it'	0.261
modal verbs of necessity	**0.428**	**proper nouns**	**0.496**	factive verbs	0.291	factive verbs	-0.208
group nouns	**0.409**	**all nouns**	**0.446**	pronoun 'it'	0.242	nominalizations	-0.218
stance nouns	**0.365**	**coord. conj. (clausal)**	**0.416**	1st person pronouns	0.200	proper nouns	-0.279
human nouns	0.303	**noun premodifiers**	**0.404**	**existence verbs**	**-0.357**	attributive adjectives	-0.310
proper nouns	0.249	**human nouns**	**0.378**	**attributive adjectives**	**-0.531**	**existence verbs**	**-0.364**
coord. conj. (phrasal)	0.202	**Indefinite articles**	**0.375**	**evaluative adjectives**	**-0.676**	**group nouns**	**-0.382**
passives	-0.221	3rd person pronouns	0.266	**abstract nouns**	**-0.757**	**concrete nouns**	**-0.400**
attributive adjectives	-0.24	evaluative adjectives	-0.216			**passives**	**-0.471**
nominalizations	-0.267	attributive adjectives	-0.249				
existence verbs	-0.286	abstract nouns	-0.291				
evaluative adjectives	-0.293	existence verbs	-0.305				
pronoun 'it'	**-0.337**	**concrete nouns**	**-0.335**				
noun premodifiers	**-0.338**	**pronoun 'it'**	**-0.396**				

first person pronouns	**-0.391**	**2nd person pronoun**	**-0.507**
place nouns	**-0.585**		