



Content list available at www.urmia.ac.ir/ijltr

*Iranian Journal
of
Language Teaching Research*



Urmia University

Suggestions toward some discourse-analytic approaches to text difficulty: with special reference to 'T-unit configuration' in the textual unfolding

Kazem Lotfipour-Saedi ^{a,*}

^a *Algonquine College, Canada*

ABSTRACT

This paper represents some suggestions towards discourse-analytic approaches for ESL/EFL education, with the focus on identifying the textual forms which can contribute to the textual difficulty. Textual difficulty / comprehensibility, rather than being purely text-based or reader-dependent, is certainly a matter of interaction between text and reader. The paper will look at some of the textual factors which can be argued to make a text more or less readable for the same reader. The main focus here will be on academic texts. The high cognitive load and low readability of the expository texts in various academic disciplines will be argued to belong to certain textual strategies as well as variations in the configurations of the T-units as the prime scaffolding for the textualization process. Different categories of these variations to be discussed here will be exemplified from a few academic and expository registers. More extensive textual analyses will, of course, be necessary in order to be able to make evidential suggestions for possible correlations between certain types and clusters of T-unit configurations on the one hand, and cognitive load and readability indices on the other, across various academic registers, genres and disciplines.

Keywords: discourse-analysis; expository texts; T-unit configuration; textual forms; textual difficulty; contextual factors; register

© Urmia University Press

ARTICLE SUMMARY

Received: 18 Oct. 2014

Revised version received: 30 Nov. 2014

Accepted: 2 Dec. 2014

Available online: 1 Jan. 2015

* Corresponding author: English Institute, Algonquine College, Ottawa, Canada
Email address: klsaedi@hotmail.com

© Urmia University Press

10.30466/ijltr.2015.20399

Introduction

Formulating our location in relation to the jargon: Register, Genre, Style, and Discourse analysis

Texts we encounter in our everyday life vary from one another in many different respects, for example, poetry, fiction prose, newspaper articles, public speeches, parliamentary question-time talks, courtroom language, classroom talk, academic papers, lab reports, billboards statistics, funeral eulogies, sermons, wedding ceremony, personal letters, business letters, degree award ceremonies, family dinner table conversation, and many others. These varieties come about thanks to variations in the choices from both lexico-grammar and para-language. But despite sharing the same resources, the variations fall under different perspectives of register, genre and style. Adopting an SFG (Systemic-functional Grammar) outlook, these terms can be defined in the following way.

Register: Any piece of text would, naturally, represent a context of situation and context of culture. Any choice made in the textualization process can be argued to represent the ethnographic features of the given context; and any, even minor, variation in these ethnographic features will be reflected in the textual choices. Register refers to this aspect of textual variation; the study of the relationship between the choices made in the textual presentation of a message and the factors involved in the ethnographic description of the related context being called register *analysis*. For example, the lexico-grammatical choices made in composing a personal letter as opposed to a business letter can be related to the relationship between the addressor and addressee and the degree of intimacy between them among many other factors.

Genre: Like register, genre is a perspective motivated by the communicative purpose and factors in the context of situation; but unlike register, it is constituted by conventional structure of the text not by lexico-grammatical variations. While a register materializes through certain choices from the lexical and grammatical resources scattered around a text, a genre comes into existence through adherence to certain conventional super-structures belonging to a whole text. For example, the way personal letters open and end; the overall structure of academic papers and the way they are segmented into sub-sections and different moves each segment can consist of all belong to generic variations. *Genre-analysis* is the study of generic variations.

Style: Like register, style perspective defines a textual orientation through lexico-grammatical choices. But while the choices in register, as noted above, are motivated by changes in the context of situation and communication goals, stylistic choices opted for are never context- motivated. They rather “reflect aesthetic preferences, associated with particular authors or historical periods.” (Biber&Conrad 2009, p. 2). For example, a rhyming pattern, a grammatical parallelism or a semantic pattern such as a metaphor employed in a literary text cannot be motivated by the context. This does not, of course, mean that the special patterns in literature-texts lack any function. For discussions on the function of the special patterns in literature cf. Leech (1970), Halliday (2002), Hasan (1985), Butt (1984) and Lotfipoursaedi (2008).

Style analysis or rather *stylistics* is the study of the special patterns mainly in literature-text.

Discourse: The terms ‘discourse’ and ‘text’ have been variously defined. ‘Discourse’ has been used to refer to ‘spoken language’, ‘language structure beyond sentence’, ‘language in use’, ‘language in action’, ‘language in context’, etc. (cf. Stubbs, 1983; Yule & Brown, 1983) and ‘text’ has been used to refer to the ‘stretch of language beyond sentence’ and ‘one or more sentences containing a whole meaning’. Some authors have used the terms ‘discourse’ and ‘text’ interchangeably (Halliday 1985). But some others use the term ‘discourse’ to refer to the ‘pre-textual process of language production’ and reserve the term ‘text’ to refer to the final ‘product’ which manifests the discourse process (cf.

Widdowson, 1984, 2004). Following Widdowson, we use 'discourse' to refer to the pre-textual process of thinking and 'text' to the surface manifestation of the discourse process.

Discourse-analysis: 'Discourse analysis' can, naturally, mean analysis of 'discourse' in any of the above senses of the term. But to be able to present a definition with the capacity to accommodate the 'ever-evolving' sense of the term 'discourse', I would like to stick to the definition formulated before (cf. Lotfipour-Saedi, 2008). Here, having presented an evolutionary history of the approaches to the study of language, and going through various approaches, starting with traditional approaches, structuralism, generativism, Chomskyan concept of linguistic competence, which excluded all the situational and contextual factors (Chomsky 1965), and the trends thereafter, which have ever since been vehemently trying to broaden the domain of this 'competence' (cf. Austin, 1962; Searle 1969; Hymes, 1972; Schegloff, 1972; Canale & Swales, 1980; Bachman, 1991), the term 'discourse' was chosen to refer to this concept of language in its ever-widening dimensions as opposed to 'language', which, prior to this socio-pragmatic surge, was used as a default term to refer to both the medium and the users' ability to use it in their everyday communication, not exceeding beyond the lexico-grammatical and phonological dimensions, of course. 'Linguistics' was the term introduced in early 20th century to refer to the study of 'language' in its lexico-grammatical and narrow sense, replacing the term 'grammar', traditionally used for this purpose. Along this line of thinking, with 'discourse' suggested to be used to refer to language in its widest dimensions, '*discourse analysis*' was used to refer to the study of language and language ability encompassing all its aspects and dimensions (cf. Lotfipour-Saedi, 2008). Discourse analysis, defined as such, would, thus, be seen as an umbrella term including text-linguistics, conversation analysis, pragmatic analysis, register analysis, genre analysis, critical discourse analysis, forensic linguistics,

Duality of language functions (core- and meta-)

Language or textual functions have been discussed by many scholars under variously labeled and almost similarly defined dual types: 'representative vs. expressive' (Buhler, 1934), 'referential vs. emotive' (Jakobson, 1960), 'descriptive vs. social' (Lyons, 1977), 'conceptual vs. interactive' (Widdowson, 1984), 'transactional vs. interactional' (Brown & Yule, 1983), and 'ideational vs. interpersonal and textual' (Halliday, 1985). In all these dual functions, one is considered as the 'core' and the other as 'supportive'. Despite their differences, since they are highly intertwined into one another, lately some authors have chosen to add the prefix 'meta-' to core function in order to refer to the supportive one: 'meta-function'. Crismore (1989), for example, has used meta-discourse to refer to the supportive discourse functions. Following Widdowson (1984), as noted above, to use 'discourse' and 'text' to refer to 'the pre-textual process' and 'the surface manifestation of that process', respectively, we have talked of discourse vs. meta-discourse and text vs. meta-text and have defined these meta-functions accordingly (cf. Lotfipoursaedi, 2005):

'Meta-discourse' is used to refer to strategies which target the discursive process of negotiating a message and which are employed to guarantee a smooth running of the discourse; and 'meta-text' is used to refer to strategies which are employed to guarantee an orderly presentation of the textual indices. (Lotfipoursaedi, 2005 p.6)

In that study, meta-discourse strategies were classified into three categories on the basis of their functions: cognitive, interpersonal and interactional, each with its own specific sub-categories; and meta-textual strategies are categorized into cohesive ties, T-unit demarcation decisions, organization announcers, paralinguistic decisions, inter-textual markers, and generic conventions. For further discussion and examples of each of these meta-strategies see Lotfipour-Saedi (2005).

In discourse-analytic approaches to language, any variation in a text, whether in the core or meta-components, is considered to be ‘motivated’; and studies under these approaches aim at making explorations in the relations between those variations and their motivating factors. The variations manifest themselves mostly in the lexico-grammar choices; and it is in this sense that grammar is seen by systemic-functional grammarians as a ‘theory of human experience, playing an essential role in construing human experience into meaning’ (cf. Halliday, 1994; Halliday & Matthiessen, 1999). Such studies can offer varying types of implications including shedding light on why the texts are the way they are, the reading and writing processes, why the same language-user may experience more / less difficulty in performing definite reading / writing tasks relative to others, and how language teacher education as well as effective language pedagogy programs can be designed addressing such issues. In the present paper, some aspects of such text variations will be examined with special reference to academic texts.

Text types and textual strategies

Texts, as noted above, represent their underlying context; and they vary in response to any variation in their contextual factors, communication goals and any element in the composition of their respective ethnography of communication. The variations in the textual forms and strategies are, thus, all motivated, and are functional in the formation of different text-types: academic, journalistic, conversational, mono-logic, dialogic, sermon, small talk, etc.

Academic texts, for example, are texts used in academia, higher educational centers, and for educational, research, argumentative and other related goals. They are mainly expository in nature, conveying information, explaining ideas, arguing for a thesis or idea and challenging another. Exposition texts are seen as “the linguistic manifestation of the genres valued in English-speaking cultures for challenging or defending the existing order of social reality.” (Chen & Foley 2004, p.190). Like any other text-types, academic texts can be of varied manifestations, depending on the goals, settings, and users involved. They can be of different disciplines like philosophy, sociology, and science. They can also vary in register, depending on who is addressed, who the address or is, and many other situational factors. For example, textbook as opposed to academic paper, classroom lecture as opposed to conference presentation, are different registers. They may further vary as genre in their organizational structure, like lab-reports, dissertations, academic papers, conference presentations, poster presentations, and classroom lectures. As noted above, generic variations are mainly rhetorical and macro-structural in nature. We said ‘mainly’ because, as discussed before (cf. Lotfipoursaedi, 2008), genre is conventional restrictions on ‘how to say’ and ‘what to say’. As an example, only certain pieces of information can be included in the ‘abstract’ of an academic paper; and the ‘preface’ and ‘introduction’ sections of a textbook will vary mostly in terms of ‘what’ rather than ‘how’.

Textual boundary, textuality, and T-unit

Textual choices, as discussed above, are all motivated by the socio-contextual factors of the given discourse production process; and it is due to such choices that one piece of academic text may be more or less readable for the same reader. The choices are mostly lexico-grammatical opted for from the relevant systems in the language. It was said ‘mostly’ because the underlying socio-contextual factors in any discourse production process will motivate not only what needs to be said but also what ‘not-to-say’. Moreover, the choices may involve those which are not lexico-grammatical, such as para-linguistic (typographic in writing and prosodic in speech), formatting and schematic structure and inter-textual (tables, figures, pictorial). A bunch of choices brought together within a communicative domain would constitute a text, which can vary in size (a single

word, a phrase, one or more sentences, one or more paragraphs, a book or a few volumes). Among the factors which bind the choices made together and contribute to the text-making process, we may name the two most important ones: socio-contextual motivation: a bunch of socio-contextual factors motivating the use of language to achieve a communication goal, and textuality: a set of text-creating agencies which help distinguish text from non-text. The textuality, as one of the essential factors, implies that the textual strategies / choices made for conveying a message are not 'dumped' in an ad hoc manner into the 'bag' of the text. They are, rather, organized into 'packages'; and what may be considered as a text on the surface is, in fact, a hierarchy of such packages (or mini-texts). One of these packages is what we refer to as T-unit.

T-unit refers to the piece of text occurring between two full-stops. It normally contains one finite verb, having superordinate relationship with any other possible finite verbs which may exist in the unit. It may also carry more than one finite verb having coordinating relationship with one another¹. A text, which represents an underlying discourse, is characterized to be made up of a set of hierarchically organized theme-rheme units (cf. Candlin & Lotfipour-Saedi, 1983). But this hierarchy should be linearly presented in the actual surface text. T-units play a crucial role in linearizing this hierarchy. In discourse-analytic studies where explorations are made as to the discursual function / value of different textual forms / choices (i.e. the way they contribute to the overall function of a text as an interface between the discourse producer and discourse receiver), the T-unit and its organizational properties in relation to its component parts play a crucial role in defining the discursual function of each of the textual choices embedded in it. To use the analogy of a political party, a T-unit acts as a political party helping to define the functional organization of all its members.

Thus, not only the type as well as the frequency of the textual forms employed within the boundary of a text but also their membership mode in its component T-units should be taken into account. For example, studying variations in the textual strategies employed in a text will certainly shed a lot of light on various aspects of the text, including its cognitive load and readability for the reader. But it is quite natural that apart from the type of a textual strategy, its 'locality' within the host text is itself an operational factor for its textual function. For example, the number of nouns opted for in the composition of a text is certainly a factor. But the functional value of nouns will vary depending on their location in the text. A noun acting as a subject and a noun placed before another noun acting as its modifier will have different values.

Decisions concerning what to be included within the boundaries of a T-unit, and all the variations in such decisions, which are motivated by the factors in the context of situation, would constitute what we call T-unit configuration.

Aspects of variations in the T-unit configuration

T-units included within the boundary of text can vary from one another in many ways: length or number of words, number of clauses and the type of relationship among them, whether superordinating or coordinating one (hypotactic or paratactic), lexical density, use of connections etc. Some of these variations will be discussed below under the title of 'aspects'.

Information structure

This refers to the way information is packaged within the textual elements and ‘what resources are available to speakers and writers for indicating to their addressees the status of information which is introduced into the discourse’ (Brown & Yule, 1983, p.153). The concept was first studied by the Prague School linguists before the Second World War and it was Halliday (1967) who brought such studies to the attention of the Western linguists. According to Halliday, information in discourse is of two types: New, that which the addresser believes is not known to the addressee, and given, that which the addresser believes is known to the addressee (either present in the context or already mentioned in the discourse). Brown and Yule (1983) maintain that to be able to take account of the different views of information status expressed by linguists on the one hand and psycholinguists on the other, a taxonomy richer than the simple ‘given / new’ distinction is needed. Using the extended taxonomy provided by Prince (1981), Brown and Yule talk about three classes of entities in a text: First class, which includes BRAND NEW and UNUSED entities; the second class, which includes INFERRABLES; and the third class, which includes EVOKED entities. By brand new, Brown and Yule (1983) mean the entities which are assumed not to be in any way known to the speaker and will typically be introduced into the discourse by an indefinite expression. They define ‘unused’ as the entities which are assumed by the speaker to be known to the hearer, to be part of his background knowledge but not in his consciousness at the time of utterance. They define ‘inferrable’ as entities which the speaker assumes the hearer can infer from a discourse entity which has already been introduced. For example, ‘the driver’ would be inferable from the interpretation of the expression the car. For Brown and Yule, the EVOKED entities can be of two types: Situationally evoked: salient in the discourse context (for instance ‘I’ and ‘you’) and textually evoked: an entity which has already been introduced into the discourse which is now being referred to for the second or subsequent time.

It sounds quite logical to assume that texts with different densities of these information type entities will vary in terms of their cognitive load. Khabbazi (1999) has carried out a cross-register study of texts in terms of the above taxonomy of information structure, finding that there is a direct relationship between comprehensibility and higher frequency of ‘situational information’.

Thematic structure

As one of the three semantic organizations of a clause, thematic structure is the one which gives the clause its “character as a message” (Halliday & Matthiessen 2004, p.64). It is through this that a clause can “contribute to the flow of discourse” (ibid, p. 64). Thematic structure in English language is achieved through placing a structural component (subject / verb / object/ complement/ adverbial) at the initial position of the clause. This initial component is called the ‘theme’, and the remainder is referred to as ‘rheme’. Theme is considered as ‘the peg’ around which the message hangs. Grammatical subject is the unmarked theme of a clause in English, the other components varying in terms of their degree of marked-ness when placed in theme position. Thematic structure was originally defined in terms of clause rank; but it can also be applied to higher ranks such as sentence, T-unit, paragraph, and even text. Thus, the element chosen to stand as the theme of a T-unit will certainly have an effect on its configuration. For example, the underlined elements in the following piece of text are the themes for their respective T-units. The choices have been motivated by the relevant contextual (discoursal) factors and as such they contribute to the natural flow of the text:

His goal, he said at the time, was a comprehensive peace agreement by the end of April that would lead to an independent Palestinian state. Despite more than 20 official talks, there is no evidence of concrete progress, but there are increasing signs that both sides may be positioning themselves to blame the other if negotiations collapse. (Editorial, New York Times Jan 3, 2014)

Other elements could, of course, have been placed in the thematic positions of these two T-units under different discoursal factors:

At the time, he said his goal was a comprehensive peace agreement by the end of April that would lead to an independent Palestinian state. *There is no evidence of concrete progress*, despite more than 20 official talks, but there are increasing signs that both sides may be positioning themselves to blame the other if negotiations collapse.

For more detailed discussion on theme-rheme, types of theme in a clause, and their definitions and examples see Halliday & Matthiessen, (2004) ; and for explorations in thematization strategies across different text-types, and also an investigation of certain dimensions of such strategies including degree of markedness, semantic and physical distance, see Lotfipoursaedi & Razaii, 1996.

Verb-frame type

Another factor affecting the configuration of a T-unit is the verb-frame (VF) type of the verbs used in the T-unit, whether as the main verb or others. Verbs in English are described to fall into many different patterns in terms of the number and types of components (subject, object, complement) they require for their syntactic-semantic operation. Many dictionaries, such as *The Oxford Advanced Learners' Dictionary* have introduced more than 40 verb patterns. We have classified all English verbs into six verb frames (VF) as follows:

1. X BE C
2. X ARRIVE
3. X EAT Y
4. X GIVE Y Z
5. X CONSIDER Y C
6. X ALLOWED Y VC

(Notes: The verbs in capital letters are examples representing a class of similar verbs. 'C' stands for 'complement' which can be a noun, an adjective, or a prepositional phrase functioning as an adverbial. 'X', 'Y' and 'Z' stand for 'subject', 'direct object' and 'indirect object'. Verbs used in VF6 fall under four types represented by verbs 'allow', 'let', 'see' and 'have', and 'VC' stands for 'verbal complement' < a verb completing the meaning of the main verb in the same VF >, which assumes the form of 'to + infinitive', 'infinitive without to', 'present participle', and 'past participle' respectively.)

For more detailed discussion of these verb frames see Lotfipour-saedi (2006).

It is natural to assume these verb-frame types to be of varying cognitive demands on the receiver. It is not, of course, claimed that VF-type is the only factor rendering some verbs more or less processable than others. Many other factors should also be considered to be operative contributing to the amount of cognitive load demanded for processing VFs. For example, the following verbs belong to the same VF (i.e. VF 1); but they can never be considered to be of the same processing difficulty for the same reader:

- (a) Mary is happy.
- (b) Mary's dreams come true.

In both of the above sentences, the verbs 'be' and 'come' belong to VF 1 {X + verb + C}; but despite this, it is obvious that the first sentence is more easily processable than the second one.

Thus, the degree of more and less readability of a sentence compared with other sentences can be attributed to the VF-type of its verb only when the other factors can be kept constant².

Verb type: Ergative vs. transitive

Verbs, apart from being categorized in terms of the process types and the number of their participants (as indicated above, see 2.3), are further classified into 'transitive' and 'ergative' models, the two making up the general system of 'transitivity' in Systemic-Functional linguistics. 'Transitives' represent the 'doing' of the action of the verb while 'ergatives' represent the 'happening' of the action. They explain the difference as 'happening' meaning that 'the actualization of the process is represented as being self-engendered' and as 'doing' meaning that 'the actualization of the process is represented as being caused by a participant that is external to the combination of Process + Medium : Agent':

The great flood spread (HAPPENING)
I'm going to send a great flood. (DOING)

These two systems are believed to be complementing one another, 'which is why they are variably foregrounded across registers: they embody different generalizations about the flux of experience, resonating with different situation types'. (Halliday & Matthysen, 2004, p. 285). Transitives are associated with 'deed - & - extension' (whether the action of the process goes beyond the Agent involving an Object) and ergatives with 'cause - & - effect' (no Agent being involved in the actualization of the process, i.e. nothing is causing it and it is being self-engendered). In fact, most of the verbs which are traditionally labeled as both transitive and intransitive in dictionaries, can have dual interpretations: transitive and ergative, the difference being in that the same participant would act as 'Actor' and 'Medium' in transitive and ergative interpretations, respectively. According to Halliday and Matthysen (2004, p.288), medium is the key participant through which the process is actualized and without which there would be no process at all: it is the entity 'through which the process comes into existence' (p.288):

Table 1

The boat	sailed
The cloth	tore
Tom's eyes	closed
The rice	cooked
My resolve	weakened
The circular muscles of the body	will contract
ACTOR (in TRANSITIVE interpretation)	PROCESS
MEDIUM (in ERGATIVE interpretation)	

In the transitive version of the above processes, the Subject of the sentences would act as ACTOR and AGENT in the transitive and ergative interpretations, respectively:

Table 2

Mary	sailed	the boat
The nail	tore	the cloth
Tom	closed	his eyes
Pat	cooked	the rice
The news	weakened	my resolve
ACTOR (in TRANSITIVE interpretation) ; AGENT (in ERGATIVE interpretation)		

Different registers as the representation of their context of situation will naturally contain different combinations of 'doing' as opposed to 'happening' verbs, with the relevant 'participants' being considered in the context of situation as ACTOR or MEDIUM.

The correct perception of such roles may be presumed to require more processing effort on the part of the reader in the reading process; and the decision for assigning the correct participant role to an entity can be argued to demand more precision in the writing process.

Physical distance between the main verb and its satellite elements

It is postulated that the process of the text comprehension proceeds upon the linear reception of its component sentences / T-units; and for each sentence to be processed, its main verb and its VF-type should be recognized. This recognition would require /necessitate that the verbal participant components (X, Y, C, Z or VC), which are referred to as 'the satellite elements' of the VF (cf. Lotfipoursaedi, 2006) should be located by the receiver (reader/listener). But because the text is linearly perceived, and due to special nature of some text-types, especially the expository texts, where elements (adverbials, modifiers and qualifiers) are added to the text which can intervene between the verb and its satellite elements, the physical distance between the verb and its satellite elements can be considered as another factor in the readability of a text. For example

Ontario Premier is threatening legal action if allegedly defamatory and false **statements made by Progressive Conservative leader Tim Hudak about her alleged involvement in a plot to delete gas-plant emails from provincial computers aren't corrected.**
(The Ottawa Citizen, March 31, 2014)

In this T-unit, the physical distance between the main verb ‘is threatening’ and its satellite elements is not so wide; but such a distance in the case of the verb in the adverbial clause, i.e. between ‘statements’ and ‘aren’t corrected’ is very wide due to a lot of intervening words (underlined above), which can lead to a delay in its processing.

Number of nouns and their semantic affiliation

Nouns in English mostly occur in word clusters called noun groups consisting of a head noun, and possibly pre-modifying (Modifiers) and post-modifying elements (Qualifiers), themselves containing possibly more embedded nouns. For a detailed description of noun group/phrase structure in English see Lotfipoursaedi (2006). Irrespective of their syntactic position, nouns, their numbers and their semantic diversity can be argued to play a role in the degree of readability of a T-unit and also a text. As an example, these features have been calculated and shown in the table 3 for the following two short pieces of texts from two different registers (i.e. newspaper editorial and fiction prose):

(Text 1)

The first change would deal with longstanding complaints from law enforcement authorities about the narrow scope of a rule stipulating that a person cannot be denied the purchase of a firearm unless he or she has been “committed to a mental institution” in the past.
(New York Times, Jan. 3, 2014)

(Text 2)

Blowing her nose, she went outside and sat on the grass to watch them play basketball.
A Courageous Battle by Suzan Bracke

Table 3

	TEXT 1	TEXT 2
# of words	45	16
# of nouns	12	3
# of nouns per words	12/45	3/16
Type/token ratio (for the semantic diversity of the nouns)	$10/12 = 0.83$	$3/3 = 1$

By the semantic diversity of the nouns contained in a text, we mean the number of different semantic fields they belong to. For this, one may choose to use the TYPE/TOKEN ratio, considering each different semantic field of the component nouns as a distinct type, and each occurrence of a noun as a token. We presume the higher the value of this ratio, the more diverse the semantic fields, which can be argued to affect the readability of the text. Normally, the value of one for this ratio means each noun belongs to a different semantic field. But when more than

one nouns contained in a text belong to the same semantic field, the denominator in this fraction will increase, lowering the value of the ratio to below one.

Variations in the cognitive load of words

Non-verb predications

The number of words in a T-unit was named as one of the factors in the cognitive load of the unit. But it is obvious that apart from the number, the type as well as the syntactic role of the words contribute to this load. Adjectives, single word (i.e. non-phrasal) adverbials, and nouns that come before other nouns to modify them are examples. These elements assume this extra load due to the invisible predication relationship they strike with the items they modify.

In other words, when an adjective is added to a noun phrase, or when a noun is used as a modifier of a head noun, they both strike an invisible predication relationship with their respective head nouns, and for this reason they can be argued to count more than single lexical items and as such they require more cognitive load (than other single words) for their processing. For example, in the following piece of text:

Two sensible changes proposed for the background check system would allow states and mental health providers more discretion than they have now in reporting information about potentially violent people. (New York Times, Jan 03, 2014)

'Sensible' represents 'changes are sensible' ; 'mental' represents 'health is mental' ; 'violent' represents 'people are violent' ; 'background' represents 'checking of background' ; 'check' represents 'system for checking' .

We can add participles (both present and past) to these categories, too, because they too perform more than a single word and would thus demand more for their processing than otherwise single words. Since in all these cases, an extra layer of meaning is 'invisibly' represented by the text, following the systemic-functional way of describing them, we may choose to call all the above cases 'grammatical metaphors. The concept of grammatical metaphor was originally explicated as the occurrence of the metaphorical realization of a meaning: semantic categories of relator, circumstance, process and quality being metaphorically realized as 'thing' (Halliday 1998).

Martin (1985) uses the term 'buried reasoning' to talk about the grammatical metaphors such as the following, where a change in the textual form hides/buries the conjunction of reasoning:

The driver drove the bus very fast, and so the brake failed.
The result of fast driving is that the brake failed.

He argues how such strategies contribute to the effectiveness of academic expository texts. One may use 'buried relations' to encompass not only the reasoning-type relations in Martin's case, but also the relations represented by adjectives, single word adverbials and modifying nouns shown above , including them all under grammatical metaphors.

Thus, any high or low frequency of occurrence of such lexical categories in a piece of text can be argued to be conducive to more or less cognitive load demanded in processing that text.

Depth of embedded-ness

Texts, as discussed before, are hierarchically organized; and as such the elements would vary in terms of their degree of embedded-ness depending on where in the hierarchy they occur. This variation, apart from other variations in the textual form, will certainly affect their readability. For example, verbs may vary not only in terms of their verb-frame type (see 2.3. above), but also in terms of whether they act as the main verb of the T-unit, or as the verb of a relative clause embedded deeper in the unit. For more concrete examples, let's look at a piece of text:

Considering the enormous value of the information he has revealed, and the abuses he has exposed, Mr. Snowden deserves better than a life of permanent exile, fear and flight. He may have committed a crime to do so, but he has done his country a great service. It is time for the United States to offer Mr. Snowden a plea bargain or some form of clemency that would allow him to return home, face at least substantially reduced punishment in light of his role as a whistle-blower, and have the hope of a life advocating for greater privacy and far stronger oversight of the runaway intelligence community. (New York Times: Editorial: 2014/01/02)

In this piece of text, the three verbs used in the first T-unit, for example, i.e. 'reveal', 'expose', and 'deserve', are not of the same degrees of embedded-ness: the first two belonging to adverbial clauses are more embedded than the last one 'deserve', which is the main verb. The elements 'at least' and 'substantially' are both adverbials; but the first one which defines the second is more embedded. Similarly, 'substantially' and 'reduced' are both modifiers; but the first one which defines the second one is more embedded. One may tend to feel that the above elements are the other way round in their degree of embedded-ness; i.e. 'substantially', being linearly prior to 'reduced', is less embedded. This can be true if 'embedded-ness' is defined in terms of the physical order. But virtually speaking, in a piece of text like: 'at least substantially reduced punishment', 'reduced' is less embedded than 'substantially' and 'substantially' is less embedded than 'at least' because they operate as follows:

Reduced (punishment)

Substantially reduced (punishment)

At least substantially reduced (punishment)

In discussing the levels of proficiency of language users and testees where they talk of the level of complexity of the structures, a language user or testee can produce or understand, one may consider this notion of degree of embedded-ness as a factor contributing to the complexity. For weighing the effect of this factor on readability, schemes may be designed to measure it numerically.

Non-finite verbs / participles

As discussed above (see 2.7.), words in a text vary in their cognitive load as a result of variations in their grammatical category as well as their locality in the text and T-unit. One aspect of the lexical

variations is for the verbs to assume non-finite (present and past participle) forms. It is obvious that, like any other variations in the textual form, such variations in the verb form are also functional and motivated. Depending on the topic and the context, only the process and not its participants are represented in the text, changing the verbs into non-finite form. This can happen in an academic and expository discourse very often where processes and not the participants are of more importance. For example, in the following piece of text, the underlined items are non-finite forms; and they are relatively abundant because the writer wants to focus on the processes under discussion and is not interested in specifying the participants. The discourse focus lies on the processes of 'editing', 'numbering' etc. not on the subjects and objects (the participants) of these actions

Sometimes separate editing efforts with mismatches between British and American edition numbering led to the existence, for many years, of two main "flavors" or "branches" of *Gray's Anatomy*: the U.S. and the British one. This can easily cause misunderstandings and confusion, especially when quoting from or trying to purchase a certain edition. For example a comparison of publishing histories shows that the American numbering kept roughly apace with the British up until the 16th editions in 1905, with the American editions either acknowledging the English edition, or simply matching the numbering in the 14th, 15th and 16th editions. Then the American numbering crept ahead, with the 17th American edition published in 1908, while the 17th British edition was published in 1909. This increased to a three-year gap for the 18th and 19th editions, leading to the 1913 publication of the *New American from the Eighteenth English*, which brought the numbering back into line. Both 20th editions were then published in the same year (1918). Thereafter, it was the British numbering that pushed ahead, with the 21st British edition in 1920, and the 21st American edition in 1924. This discrepancy continued to increase, so that the 30th British edition was published in 1949, while the 30th and last American edition was published in 1984. (http://en.wikipedia.org/wiki/Gray's_Anatomy)

In terms of the cognitive load, non-finite verbs can be argued to require more cognitive effort than their parallel finite forms. In our discussion of the lexical forms such as adjectives, one-word adverbials, and nouns acting as modifiers (when placed before other nouns), we argued how such forms are in fact functioning as predicates with deleted participants (see 2.7. above); and following Martin (1985), who uses the term 'buried reasoning' to refer to cases of grammatical metaphor where elements of intended meaning are missing (hidden / buried) from the surface text, we labelled such lexical forms as 'buried forms' because their actual discursive 'value' is different from their syntactic 'signification' (if we may borrow the binary distinction of 'signification' vs. 'value' from Widdowson, 1978). I would now like to expand the domain of the term 'buried form' to include non-finite verbs and in fact, all grammatical metaphors (nominalizations).

Noun phrase and its size

Nouns mostly occur in the form of phrases consisting of two or more words. Extra words added to the head noun perform various grammatical and cognitive functions. Grammatical because they are either added before the head noun (Modifiers) or after the head noun (Qualifiers) adding more information on it. Cognitively and discursively because they are seen to be cognitively and discursively motivated by the ethnographic composition of the context of situation. Those added before the head noun (modifiers) fall into five categories (of determiners, adjectives, nouns, and participles: both past and present) and those added after the head noun (qualifiers) are of two categories of 'of + noun phrase' and 'relative clauses'. These qualifying relative clauses may be shortened in different ways (see Lotfipoursaedi, 2007 for a detailed discussion of modifiers and qualifiers: their types and variations).

The type and number of defining elements added to the head noun in the noun phrases of a T-unit in a text are naturally discursively motivated, i.e. the variations respond to the factors in the context of situation as they are perceived by the discourse producer (speaker / writer). Variations in their types and numbers in a text can be argued both to be reflective of factors in their respective context of situation and to have effects in the readability of the text.

Adverbial clauses

These are clauses added to the body of a T-unit defining different aspects of the performance of the main verb in the unit. For example, in the following T-unit,

But when Mr. Abbas welcomed the latest group to the West Bank this week, Mr. Netanyahu accused him of embracing terrorists, even though Mr. Abbas never condoned the prisoners' crimes. (*New York Times* 03 Jan 2014)

there are two 'adverbial clauses': '*Mr. Abbas welcomed the latest group to the West Bank this week*' and '*Mr. Abbas never condoned the prisoners' crimes*' both defining the T-unit (underlined above). They are connected to the body of the unit by discourse markers/connectors 'when' and 'though'. Many variations can occur in the configuration of a T-unit in relation to the adverbial clauses: the number of the adverbial clauses, whether or not they are short-form or complete, and their location in the T-unit (preceding or following the main clause), the presence or absence of connectors, and the degree of marked-ness of the connectors³. Such variations, like any other variations in a text, are all reflective of the factors in the relevant context of situation differing across genres, registers and disciplines and will naturally have effects on the cognitive load of texts.

For a detailed discussion of the adverbial clauses, their taxonomies in terms of the meaning spectra they define, the way they are shortened, and the taxonomies of connectors / discourse markers see Lotfipour-Saedi (2006).

Verbal Modulators

Verbs are almost always used with what we have chosen to call 'verbal modulators' (VM). These elements represent factors in the context of situation and may be textually represented through various lexico-grammatical devices such as 'mood, first verb, reporting constructions, emphasis, negative, question, fronting, extraposition, passive and tag-questions' (Lotfipour-Saedi, 2006, p. 98). In the following piece of text, the underlined elements are verbal modulators:

The framework must not become yet another interim agreement that leaves Palestinians with empty promises. To succeed, it will need to be embraced and defended by Mr. Netanyahu and Mr. Abbas, who must acknowledge that neither society will be secure until both can learn to compromise and live as states, side by side. (*New York Times* 03 Jan 2014)

Verbal modulators would encompass various functions discussed under different names: for example, 'stance', 'and hedging', 'epistemic modality' (Coates, 1983), 'evaluation' and 'appraisal' (Matin & White 2005), and 'interpersonal meta-discourse' (see Lotfipour-Saedi, 2005).

For a detailed discussion of the taxonomies and functions of different verbal modulators and their textual manifestations in English see Lotfipour-Saedi (2006).

T-unit configuration vs. textual unfolding

As discussed above (section 2), T-units may vary in their configuration. Apart from the aspects discussed above, one may also want to look at the mode of such variations in the component T-units within the boundary of a text. In other words, we are not interested in the variations of T-unit configuration per se but rather in such variations among the T-units of a text as it unfolds. Discourse analytic studies should be interested in investigating how the T-units contained within the boundary of a text vary from one another and in relation to one another in terms of their configuration and would consider all such variations as socio-cognitively motivated. One can look at inter-T-unit variations across various registers, genres and disciplines, variations such as the length of the T-units, the thematization and thematic development across the T-units, lexical density, semantic relation between T-units, use of discourse markers (conjunctions), degree of cohesiveness, contrastively over all the T-units in a text studying the changes in relation to the unfolding of the text. For example, in terms of the T-unit length, having calculated the average number of words per T-unit in a text (by dividing the total number of words by the number of T-units in the text), one may see how the length of each T-unit varies in relation to this average value. It is, of course, natural for the T-units of a text to contain unequal number of words. But spiky variations can be meaningful. Examples of texts can be found in any register from different disciplines where very long T-units are punctuated by considerably short units. One can examine such examples in terms of the units preceding or following them and their meaning or communicative function, as well as the locality of the segment / paragraph in which they are situated in relation to the whole text or in terms of the generic structure of the text, for example, in abstracts as opposed to the introduction or conclusion sections of paper. As an example, in the following piece of text from a newspaper editorial article where the average number of words per T-unit is 25.1, the last T-unit contains only 8 words:

The first would deal with longstanding complaints from law enforcement authorities about the narrow scope of a rule stipulating that a person cannot be denied the purchase of a firearm unless he or she has been “committed to a mental institution” in the past. This ignores whole categories of obviously risky citizens (New York Times Jan 03, 2014).

And in the next piece of text from another editorial where the average number of words per T-unit is 27.3, the first T-unit contains only 5 words:

Signs of failure are everywhere. On Thursday, standing beside Mr. Kerry, Israeli Prime Minister Benjamin Netanyahu delivered a harsh assessment of his Palestinian counterpart, President Mahmoud Abbas, and, implicitly, the prospect of a Middle East peace agreement. Days earlier, Israel let it be known that it would build more settlements in the West Bank, further poisoning the political atmosphere while shrinking the territorial space for a deal. Hard-liners in Mr. Netanyahu’s government are pushing a bill that would annex settlements in the Jordan Valley area of the West Bank, where about 6,000 Israeli settlers and 10 times as many Palestinians live. (New York Time: Jan 03, 2014)

Speculations can be made about the variations in all the above-named parameters in the T-units of texts (length, thematization strategies,) and the possible underlying factors motivating them (including, the topic, intended audience, the producer’s state of mind, genres and generic moves,) and any consistent conclusions arrived at can offer insights about reading and writing processes and their pedagogy.

Implications

As indicated in the title, this paper includes a set of suggestions towards the characterization of the notion of textual difficulty with special reference to the T-unit configuration in the textual unfolding. The claims made about the changes in the configuration of the T-units in a text and their effect on the cognitive load demanded for its processing are based on the author's theoretical postulations and experiential intuition as an ESL/EFL educator and are yet to be tested experimentally. For this reason, one cannot talk of the actual implications of the study. However, potential implications can be cited for comprehensive studies, including the textual analysis along the methods suggested in this paper together with the experimental support for the effect of a single or a specific bunch of textual configurations on the textual difficulty and readability.

It is obvious that studies along the above-discussed suggestions on the nature of expository texts and the variations they exhibit in terms of the number and type of discourses and textual strategies across different registers, genres and disciplines will offer plenty of awareness to the professionals in ESL / EFL education, enabling them to determine their course objectives, design their curricula, decide on their teaching and classroom tasks, choose the right texts suitable for their target clients (both in terms of text –type and the clients' level of proficiency) for their materials as well as their tests, and plan for their testing tasks.

References

- Austin, J.L. (1962). *How to do things with words*. Oxford: OUP.
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: OUP.
- Biber, D. & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: CUP.
- Brown, G. & Yule, G. (1983). *Discourse Analysis*. Cambridge: CUP.
- Buhler, K. (1934/1965). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Stuttgart: Gustav Fischer.
- Butt, D. (1984). 'To be without a description of to be': *The relationship between theme and lexico-grammar in the poetry of Wallace Stevens*. Unpublished Ph.D. Thesis. Macquarie University, Sydney.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-48.
- Candlin, C.N. & Lotfipour-Saedi, K. (1983). Processes of discourse. *Journal of Applied Language Study* 1(2), 103-133.
- Chen, Y. & Foley, J.A. (2004). Problems in the metaphoric reconstrual of meaning in Chinese EFL learners' exposition. Revelli, L.J. & R.A. Ellis, (eds.), *Analyzing Academic Writing* (pp. 109-209). London: Continuum.
- Chomsky, N. (1965). *Aspects of theory of syntax*. Cambridge, MA.: MIT Press
- Coates, J. (1983). *Semantics of the modal auxiliaries*. London: Croom Helm.
- Crismore, A. (1989). *Talking with Readers: Metadiscourse as Rhetorical Act*. Indiana University Press.

- Halliday, M.A.K. (1967). Notes on transitivity and theme in English. Part 2. *Journal of Linguistics*, 3, 199 – 244.
- Halliday, M.A.K. (1985). *An introduction to functional grammar*. London: Arnolds.
- Halliday M.A.K. (1998). Things and relations: grammaticising experience as technical knowledge. Martin, J.R. & Veel, R. (eds.), *Reading Science*. London: Routledge.
- Halliday, M.A.K. (2002). Linguistic function and literary style: An inquiry into the language of William Golding's *Inheritors*. Webster, J. (ed.) *Linguistic Studies of Text & Discourse*. London: Continuum.
- Halliday, M.A.K. & Matthysen, C.M.I.M. (1999). *Construing experience through meaning: A language-based approach to cognition*. London: Continuum.
- Halliday, M.A.K. & Matthysen M.I.M. (2004). *An Introduction to Functional Grammar*. London: Arnolds.
- Hasan, R. (1985). *Linguistics, language and verbal art*. Victoria: Deakin University Press.
- Hymes, D. (1972). On communicative competence. Pride, J.B. & Holes, J. (eds.), *Sociolinguistics: Selected readings* (pp. 269-293). Harmondsworth, England: Penguin.
- Jakobson, R. (1960). Closing statement: Linguistics & poetics. Sebeok, T. (ed), *Style in Language*. (pp. 350 -77). Cambridge, Mass.: MIT Press
- Khabbazi, L. (1999). *Explorations in the role of information structure on readability*. MA dissertation, University of Tabriz, Iran.
- Leech, G. N. (1971). "This bread I break" –Language and Interpretation. In Freeman, D. C. (1970) *Linguistics and Literary Style* (pp. 119-128). New York: Holt, Rinehart & Winston Inc.
- Lotfipour-Saedi, K. (2005). *Explorations in meta-discoursal and meta-textual strategies: Implications for ESL writing pedagogy*. A paper presented at TESL Canada Conference, May 26-28, 2005, Ottawa, Canada.
- Lotfipour-Saedi, K. (2006). *Towards the Textuality of a Text: A Grammar for communication*. Tabriz: Forouzesh.
- Lotfipour-Saedi, K. (2008). *Discourse and the Act of Translating: A discoursal approach to the characterization of translation equivalence*. Tabriz: Islamic Azad University Press.
- Lotfipour-Saedi, K. & Rezaii, F. (1996). Explorations in thematization strategies and their discoursal motivation. *Text*, 16(2), 225-249.
- Lyons, J. (1977). *Semantics* (Vols. 1 & 2). Cambridge: CUP.
- Martin, J. R. (1985). *Factual writing: Exploring and challenging social reality*. Geelong, Vic: Deakin University Press.

- Martin, J. R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.
- Prince, E.F. (1981). Towards a taxonomy of given-new information. P. Cole (ed.) *Radical Pragmatics* (pp. 223-256). New York: Academic Press.
- Searle, J.R. (1969). *Speech Acts*. Cambridge: CUP
- Stubbs, M. (1983). *Discourse Analysis: The sociolinguistic analysis of natural language*. Chicago: University of Chicago Press.
- Widdowson, H. G. (1978). *Teaching English as communication*. Oxford: OUP.
- Widdowson, H.G. (1984). *Explorations in applied linguistics 2*. Oxford: OUP.
- Widdowson, H. G. (2004). *Text, context, pretext: Critical issues in discourse analysis*. Oxford: Blackwell Publishing.

Kazem Lotfipour-Saedi is a professor of applied linguistics and has taught in universities in Iran, Australia, Japan and Canada. He has authored papers and books in discourse analysis, ESL/EFL methodology, and translation studies.

¹ Although the T-unit concept here and the SFG concept of 'clause complex' may seem to be identical, we have chosen to use T-unit, and not 'clause-complex', to highlight its function in textual unfolding.

² We are mainly looking at the textual strategies and forms here. Otherwise, texts can vary in terms of their cognitive load and readability in terms of many non-textual factors such as whether the nouns and verb denote concrete or abstract object or actions, how common and everyday these objects and actions are for the reader, etc.

³ By degree of marked-ness of the connectors, we mean the degree to which a connector is not prevalent in everyday use. For example, among the connectors used for the adversative relationship, the most and least prevalent connectors can be said to be 'but' and 'notwithstanding the fact that' respectively; and among the connectors used for the additive relationship, the most and least prevalent connectors may be said to be 'and' and 'moreover' respectively. These examples are, of course, based on our intuitive knowledge of the everyday language use. It is obvious that studies need to be conducted for a more documented use of the connectors and their degree of marked-ness across different genres, registers and disciplines.