

## **A Systemic Dynamics Model of Text Production**

Giacomo Figueredo

*Department of Language Studies, Federal University of Ouro Preto, Mariana, Brazil*

Address: Departamento de Letras – DELET/ICHS/UFOP. Rua do Seminário, s/n –  
Centro – CEP: 35420-000 – Mariana/MG – Brasil.

Email: giacomojakob@gmail.com

Grazziela P Figueredo

*The Advanced Data Analysis Centre, The School of Computer Science, The University  
of Nottingham.*

Address: Wollaton Road, Jubilee Campus, Nottingham, UK, NG8 1BB.

Email: Grazziela.figueredo@nottingham.ac.uk

Biographical notes:

Giacomo Figueredo is a Senior Lecturer of Linguistics at the Federal University of Ouro Preto, Brazil. His research interests include quantitative methods in linguistics, systemic theory and language dynamic modelling. His research focus has been on language probabilistic descriptions and multilingual text production, focusing on exact descriptions of systemic probability and dynamics across text types in different languages.

Grazziela P. Figueredo is a Senior Research Data Scientist at the Advanced Data Analysis Centre at The University of Nottingham. The focus of her research is the development and application of techniques for systems simulation and intelligent data analysis in big data sets. She has been working with data analysis and simulation methods for a wide range of areas, including academic, medical and industrial partners. Her research interests include artificial immune systems, knowledge discovery in databases, bio-inspired computing and multi-scale modelling and simulation.

## **A systemic dynamics Model of Text Production**

**Abstract:** This paper introduces a quantitative model of text as it unfolds in time. The model conceptualizes text as a functional unit of language. This organization can be difficult to identify because it forms complex patterns of linguistic laws, probability and dynamics. These patterns are covert configurations and need complex methods to be investigated. One such method is to draw from qualitative frameworks derived from the quantitative properties of language. Previous studies (Plum & Cowling, 1987; Rybicki & Eder, 2011; Zhang & Liu, 2017) have shown that covert configurations can be obtained through qualitative frameworks. When dynamics is considered, however, a model of text production including the variable time is needed. This paper therefore aims at addressing this research gap by proposing a dynamics model of text unfolding. It draws from systemic theory and models its categories quantitatively. Time is introduced as variation of choice. The model is applied to a sample of text. Results show how individual choices contribute to text unfolding – describing the amount of meanings at any given moment in text time. In addition, the dynamic accumulation indicates core characteristics of a text, which can be further explored in text behavior and typology.

**Keywords:** systemic dynamics; text modelling; system networks; valeur; agnation

### **Introduction**

As an ever evolving discipline, Quantitative Linguistics has increasingly informed other disciplines and applications in a number of correlated areas, such as natural language processing (Dione, 2014), stylistics (Rybicki & Eder, 2011) and translation (Ke, 2012; Rybicki & Heydel, 2013). Consequently, Quantitative Linguistics meets ever growing challenges, “as the number of fields of research increases, the number of problems increases as well” (Tuzzi, Benešová, Mačutek, 2015, p. v).

One such challenge is to further develop quantitative methods to explain language organization as a system, in particular for phenomena which are not merely typical – expressed by overt structuring such as phonemes, morphemes, particles or words; but are rather cryptotypical configurations (Whorf, 1956) expressed by linguistic laws, patterns of choice in systems, or relations between system and structure.

A comprehensive account of quantitative properties of language grounded on overt structuring is found on the literature (e.g. Popescu, Mačutek, Altmann, 2009; Altmann, 2015) – whose contributions are undeniably fundamental. However, the fact that their methodology focuses on language properties realized by "visible" overt structures leads to the problem of leaving properties realized by "invisible" patterns and relations unaccounted for.

Attempts to explain quantitative properties of cryptotypical configurations, particularly those generated by covert patterning of systems, also figure in quantitative work (e.g., Plum & Cowling, 1987; Rybicki & Eder, 2011; Zhang & Liu, 2017). However, their methodology focuses on properties of language from a static perspective – that is, such methods models language product (the text) as a static frame where categories occupy a pre-determined locus in the structure. The fact that these studies focus on static modelling leads to the problem that they may be missing out on the dynamic process<sup>1</sup> of language unfolding as text.

Due to its empirical nature, both typical and cryptotypical quantitative studies naturally provided limited insight into language functionality as a whole. However, we believe that a possible integration between quantitative descriptions and the functionality performed by the categories under investigation in the process of

---

<sup>1</sup> In this paper, what we mean by 'language as process' is a model of successive choices in language systems leading up to the production of a text; it is not related to the cognitive processes of an individual speaker/writer. In this paper we model the final product of text production and for the purposes of this paper we are not modelling the several re-writings, corrections and tentative versions of the text leading up to the final product.

deploying language to make meaning could broaden the scope of research in quantitative language modelling.

Given the opportunity to further the understanding of the quantitative nature of language, we approach quantitative language investigation from a systemic and functional perspective. More specifically, we address the following problems – (i) investigating cryptotypical configurations quantitatively, while (ii) capturing the dynamics of their unfolding in text and (iii) starting an integration between quantitative properties and systemic functionality. As a result, the paper has the objective of introducing a quantitative model to describe the meaning-making process (Halliday, 1996; Lemke, 1984, 1991) of language systems and their impact in the production of text as it unfolds.

The paper departs from language systems organization to make meaning that is functional to its context (Halliday & Hasan, 1976; Halliday, 1978; Martin, 1992; Lemke, 1993). It then develops a method to quantifying systemic properties of language – which enables the calculation of the different "work" language categories performs. Finally, it introduces the variable time to the model, enabling the description of the meaning-making process at any given moment in a text. Since time is change in space, the model sets up a language space derived from measurements of language categories, and so text is not seen as a static structure anymore, but rather as unfolding movement in the language space.

The model can contribute to quantitative studies by broadening the scope of research to quantifying cryptotypical configurations and measuring language functionality dynamically. In addition, the model has the potential to provide insight into language explanations on meaning-making based on quantitative properties of

language, particularly when text production is involved, such as in language generation, education, translation and text typology.

## **1. Theoretical underpinning**

### ***1.1. Less traditional measures in Quantitative Linguistics***

Studies at the core of Quantitative Linguistics (Tuzzi, Benešová, Mačutek, 2015) tackle the challenge of developing quantitative methods to explain language organization is carried out by solid, empirical methods into examining cryptotypes as laws governing language behavior such as distribution, frequency, length, probability and dynamics. To this end, core Quantitative Linguistics focuses on categories which have an overt realizational structure, associated with a corresponding graphological/phonological expression, such as studies on words and morphemes (Popescu, Mačutek, Altmann, 2009) or overt elements of text based on parts of speech (Naumann, 2015). Text itself is approached as a configuration of cryptotypical distribution expressed by overt structures, or “a linear sequence of meaningful entities, organized also hierarchically (e.g. in the hierarchy sentence, clause, phrase, word, morpheme, syllable, phoneme)” (Popescu *et al.* 2015, p. 2).

Some studies in Quantitative Linguistics, particularly those which are not at the core of the discipline, have attempted to meet the challenge by applying “less traditional measures to scrutinize diversity of language and its properties” (Tuzzi, Benešová, Mačutek, 2015, p. v-vi), for which they have relied to some extent in language categorization of covert configurations derived from qualitative descriptions as a means to understand language quantitative properties (Plum & Cowling, 1987; Rybicki &

Eder, 2011; Zhang & Liu, 2017). These studies show that cryptotypical configurations, such as genres or rhetorical relations can also be studied by exact methods, despite the fact they are not realized by any one overt structure, but by patterns, relations and probabilities of relations.

One such area of less traditional measures that has systematically drawn methodologies and results from quantitative studies is the functional linguistics dedicated to language's organization as system – known as systemic functional linguistics (Halliday & Kress, 1976).

Systemic functional linguistics (SFL) has always benefited from quantitative methods, placing frequency distribution, probability and patterning at the center of system's organization (Nesbitt & Plum, 1988; Halliday & James, 1993; Matthiessen, 1999, Halliday, 2005). In their seminal paper Nesbitt and Plum (1988, p. 6) assert that “since language itself consists of patterns which can only be quantitatively delineated, studies of [quantitative] type are central to the core questions of linguistics and important for many applications of linguistic science”.

Further expounding on the notion of system will be seen below.

## ***1.2 Language as system***

SFL (Halliday 2002) views language as a semiotic system organized as choice between features in a set. Each feature acquires meaning against the background of choices for other features in the set that could have been made but were not. For any given language subsystem, the job done by each language stratum - phonology, grammar and semantics - is to change (responding to the pressure of new contextual demands) systemic (paradigmatic) organization of features in order to create meaning.

Whenever there is need for a reshaping of some aspect of human life there is also a contextual pressure for new meanings "packed" as texts. Language strata reorganize features of systems, changing both their paradigmatic contrast and their probability, thus creating new meanings through variation and new texts as patterns of frequent variation.

As a result, language is understood in terms of (Halliday 1991) (i) its relations to the context of culture – the “environment” in which it is meaningful; and (ii) the process in which language as a reservoir of meaning-making potential (the system) becomes, through operations in phonology, grammar and semantics, language in context (text).

As a meaning-making system, language organizes itself (see section 1.2.1 below) through relations among systems, language patterns frequencies of choice into probable co-selections. This “meta-organization” of language in relation to its internal functions is known as metafunctions (Halliday 1978). Co-selections of features from different systems lead up to three metafunctions (see section 2.6 below) – interpersonal, ideational and textual (Halliday and Matthiessen 2004).

If we focus on the clause grammar stratum - the unit of analysis of this paper - we can see that in the interpersonal metafunction, the grammar systems of mood, modality and polarity enact social interaction through clause types: indicative (declarative/interrogative) and imperative; and evaluation and assessment. Modality enables assessing propositions according to the degrees of modalization (probability and frequency) or modulation (obligation and inclination). Polarity is realized grammatically by absolute degrees of commitment between speaker and proposition

In the ideational metafunction, transitivity is the grammar system assigned to represent things and events as grammatical functions of Participant and Process, respectively. Transitive representations can be typologized in a general form as Material

(representation of the events external to Participants), Mental (internal events to Participants – or consciousness), Relational (relations between Participants), Verbal (symbolic events). Logical relations chain up experiential meanings in sequences of adding (extension), restating (elaboration), focusing on specific aspects (enhancing). Projection sets up a semiotic reality (realis/irrealis) in terms of ideas, desires, wishes, sayings, hypotheses, etc.

In the textual metafunction, the main textual grammar system is Theme.

Discursively, Theme rearranges each clause to fit context within text. Its main function is either to keep the arrangement of the discourse flow or shift the arrangement to best fit contextual/text types new phases. In addition, it enables interpersonal and ideational meanings by contextualizing them in a specific situation, according to a specific text type (see examples in Figure 1).

clause example 1						
metafunction	system	feature (choice)	function in structure			
			<i>you</i>	<i>forced</i>	<i>the air molecules</i>	<i>inside the balloon</i>
interpersonal	MOOD	declarative	Subject	Finite/Predicator		
	POLARITY	positive				
	MODALITY	non-selected				
ideational	TRANSITIVITY	material	Actor	Process	Goal	
	CIRCUMSTANCE	location				Circumstance
textual	THEME	theme-subj.	Unmarked			
	PREDICATION	non-selected				
clause example 2						
metafunction	system	feature (choice)	function in structure			
			<i>your ears</i>	<i>can</i>	<i>hear</i>	<i>the noise loudly</i>
interpersonal	MOOD	declarative	Subject	Finite	Predicator	
	POLARITY	positive				
	MODALITY	modalization		Probability		
ideational	TRANSITIVITY	perception	Senser	Process		Phenomenon
	CIRCUMSTANCE	manner				Circumstance
textual	THEME	theme-subj.	Unmarked			
	PREDICATION	non-selected				



Figure 1. Grammar systems analysis of the clause.

### 1.2.1 System modeling

The systemics dynamics introduced in this paper is a quantitative model of the meaning-making process based on SFL (Halliday, 1996; Martin, 2013), where axial relations (choices in paradigm realized by structures in syntagm) are foregrounded. Accordingly, it develops its account of language primarily as a system.

A system comprises a given set of signs related through the principle of valeur (Saussure, 1966). Signs have a relative value in relation to one another because they (i) sustain a degree of similarity, yet (ii) they are different (Figure 1).

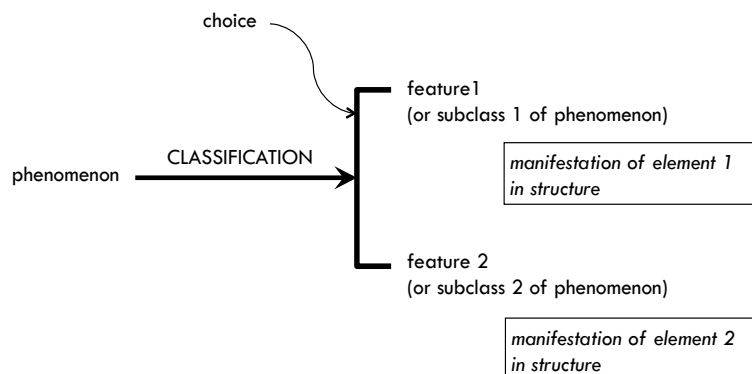


Figure 1. A system.

From left to right, a system begins with an entry condition, or the unit of phenomenon being classified by the system. Then, the arrow and angle bracket together show the classification criterion; they indicate features and how many subclasses of the phenomenon there are. The boxes under features show their manifestation in structure – i.e., their realization.

All phenomena in language can be represented paradigmatically as a system. For example, the system of VOICE<sup>2</sup> in the verbal group of English is described as shown in Figure 2.

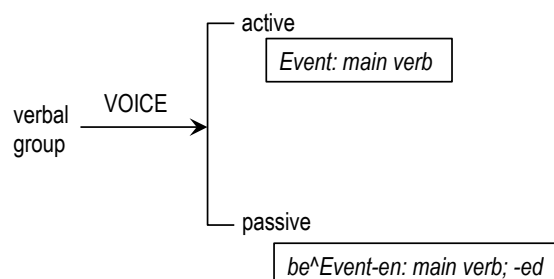


Figure 2. The system of VOICE in English.

Verbal group is the entry condition. ‘Active’ and ‘passive’ are features. ‘Active’ is realized in structure by main verbs with the function of Event; ‘passive’ is realized by the verb ‘be’ followed by a main verb functioning as Event and morphemes ‘-en’ or ‘-ed’.

### 1.3 Cryptotypes

Features of systems can be realized by overt structures (Halliday & Webster, 2005) such as elements (nouns, verbs, morphemes, phrases, etc.), function-structures (Subject, Modality, Auxiliary, Event, etc.), ordering (determiner followed by noun). These are categorized as typical because they can be “seen”.

Cryptotypes however are realized by covert structures – or a group of reactances (Whorf, 1956) that have no apparent manifestation in structure, thus cannot be “seen”. The only way cryptotypes can be found is by probing language.

---

<sup>2</sup> All systems names are conventionally written in small caps; system of VOICE, system of MOOD, etc.

An example of cryptotypes is the grammar for “mass nouns” in English. Mass nouns (e.g. water) have no apparent manifestation – i.e., there are no morphemes, words, prosodies or sequences “showing” uncountability. “Mass” is cryptotypical, being realized by a relation of selections in several systems. If ‘water’ is used as an example to be probed: (a) a cardinal numerative cannot be added, as in “one water is made up of tiny particles”. (b) A capacity can be added, as in “a liter of”. (c) A plural form can be used, but it will not mean “more than one”; but rather “different types”, as in “international waters” (Figure 3).

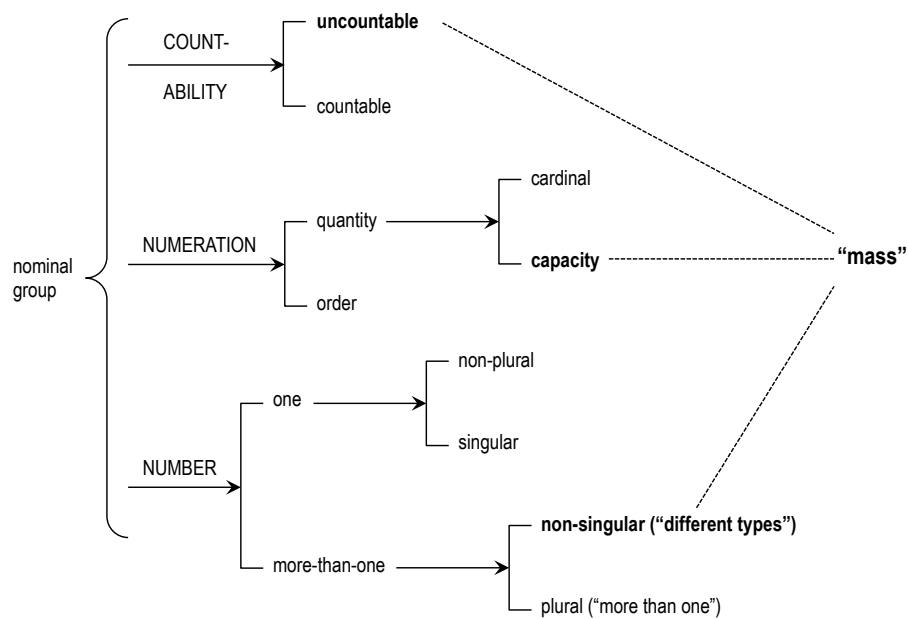


Figure 3. Cryptotypical realization of “mass” in English.

Because of its cryptotypical organization, language can be described as a vast network of systems. Systems can cosseselect and form patterns. A choice in a given system is conditioned by choices in other systems (Matthiessen, 1995).

### 1.4 Valeur: agnation and delicacy

Valeur involves contrasts of features – i.e., agnation; and levels of subsequent choice – i.e., delicacy. Agnation (Gleason, 1965) is the factor of valeur operating contrast. Delicacy is the factor of valeur operating detail (Halliday, McIntosh & Stevens, 1964). Every system has an entry condition and (at least) two contrasting features. Features, in turn, become entry conditions to subsystems in higher levels of delicacy. Figure 4 shows this process in MOOD: IMPERATIVE in French (Caffarel, 2004, p. 93).

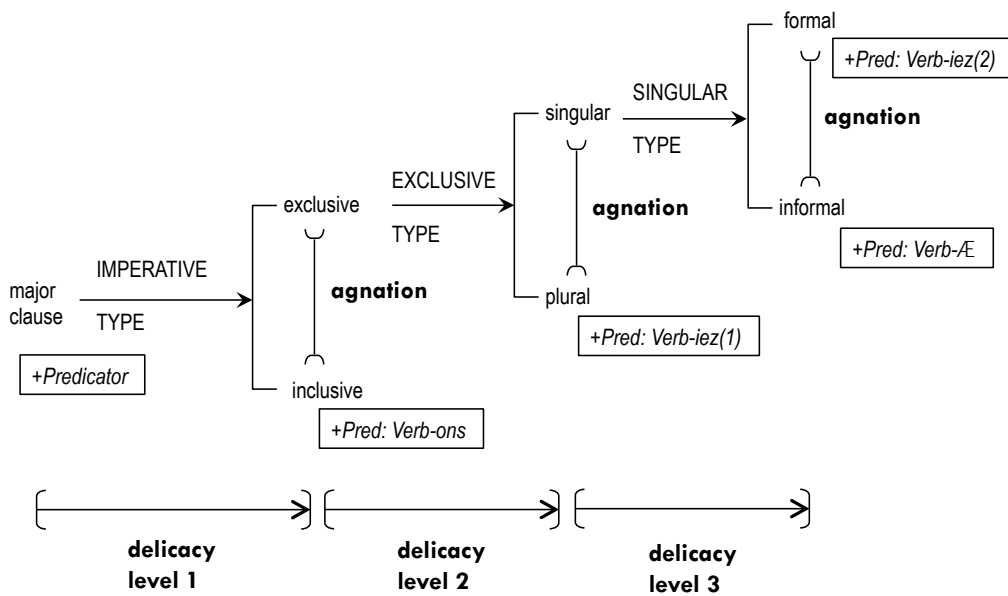


Figure 4. Agnation and delicacy.

MOOD: IMPERATIVE in French has 3 subsystems (IMPERATIVE, EXCLUSIVE and SINGULAR), respectively with one agnation each (inclusive\exclusive; singular\plural; formal\informal). It has 3 levels of delicacy (from major clause to inclusive\exclusive; then to singular\plural; then to formal\informal).

The qualitative modelling of language paradigmatic organization presented above will be quantified below in Section 2, so it can be used to measure meaning-making.

## **2. Methods: systemic properties of language**

In this section we introduce the mathematical formulation enabling the methods for measuring meaning-making in language. We begin by quantifying valeur and using this quantification to establish scales that will compose the language space where texts unfold.

Valeur is a principle of any systems and it can be used as a measure of meaning, that is, the demand for a greater amount of meaning to be made implies system networks with a greater number of valeurs. The systemic dynamics model introduced in this paper is conceived as a device to quantify the number of valeurs produced by systems. When the valeurs of systems are generated considering the variable time, then, accordingly, a quantitative model of the meaning-making process is obtained.

### ***2.1 Quantifying agnation and its properties***

The two factors determining agnation (represented by ‘ $a$ ’) are the number of features ( $fe$ ) and the contrast ( $c$ ) between them. The contrast is calculated by combination of features 2 at a time. The value of  $c$  is calculated as follows:

$$c_2^{fe} = \frac{fe!}{2! \times (fe-2)!} \quad (1)$$

If a system had less than 2 features the combination calculation results in the factorial of a negative number. Therefore a size of features smaller than 2 would not make any meaning, because there would not be a defined value for contrasts, as shown in equations 2 and 3. As a result, there must be at least 2 features at all times.

$$c_2^0 = \frac{0!}{2! \times (0-2)!} = \frac{1}{Undefined} \quad (2)$$

$$c_2^1 = \frac{1!}{2! \times (1-2)!} = \frac{1}{Undefined} \quad (3)$$

Features and contrasts are both directly proportional to agnation, and agnation can be described as the product of features and contrasts in a given system:  $a \propto fe$ , if  $c$  is constant;  $a \propto c$ , if  $fe$  is constant. The equation defining the agnation  $a$  calculation is:

$$a = fe \times c \quad (4)$$

For example, VOICE in English has 2 features, ‘active’ and ‘passive’. Thus, agnation equals 2 (as  $c = 1$ ). TENSE in French has 3 features, ‘past’, ‘present’, ‘future’; thus agnation equals 9 ( $c = 3$ ). THEME in German has 4 features, ‘finite’, ‘textual’, ‘interpersonal’, ‘experiential’; agnation therefore equals 24 – as shown in equations 5, 6 and 7:

$$a_{voice} = fe_{voice} \times c_2^{fe_{voice}} = 2 \times c_2^2 = 2 \times \left( \frac{2!}{2! \times (2-2)!} \right) = 2 \quad (5)$$

$$a_{tense} = fe_{tense} \times c_2^{fe_{tense}} = 3 \times c_2^3 = 3 \times \left(\frac{3!}{2! \times (3-2)!}\right) = 9 \quad (6)$$

$$a_{theme} = fe_{theme} \times c_2^{fe_{theme}} = 4 \times c_2^4 = 4 \times \left(\frac{4!}{2! \times (4-2)!}\right) = 24 \quad (7)$$

Although agnation can grow indefinitely in principle, this is not what is found in data extracted from actual language descriptions. An important quantitative property of agnation is that most systems in natural languages have from 2 to 4 features. As an example, Table 1 shows the number of features per system described in Caffarel *et al.* (2004) – a systemic description of grammar in French, German, Japanese, Tagalog, Chinese, Vietnamese, Telugu and Pitjantjatjara.

2 features	3 features	4 features	5 features	6 features	more than 6 features	Total
340	36	15	3	1	0	395

Table 1. Features per system in Caffarel *et al.* (2004).

Every increase in agnation means an increase in the meaning-making potential of a system. Agnation grows incrementally relative to a previous amount. The rate between subsequent amounts of agnation shows the amount to be gained when new features and contrasts are introduced to a system. Table 2 shows in the last the rate between subsequent levels of agnation ( $a_i$  and  $a_{i+1}$ ).

features	contrasts	agnation $a_i$	agnation $a_{i+1}$	rate $(a_{i+1})/a_i$
1	0	0	2	----
2	1	2	9	4.50
3	3	9	24	2.66
4	6	24	50	2.08
5	10	50	90	1.80

6	15	90	147	1.63
7	21	147	224	1.52
8	28	224	324	1.44
9	36	324	450	1.38
10	45	450	605	1.34

Table 2. Agnation growth rate.

Even though agnation increases as more features are added, its incremental growth rate relative to features and contrasts reduces asymptotically, with a limit tending<sup>3</sup> to 1, as shown in Figure 5. That is, if  $a$  is considered as a function of  $fe$ , ( $a(fe)$ ) then:

$$\lim_{fe \rightarrow \infty} \frac{a(fe+1)}{a(fe)} = 1 \quad (8)$$

5-feature systems seem to be the point at which most systems stop adding features, since the payoff in meaning-making is steadily decreasing.

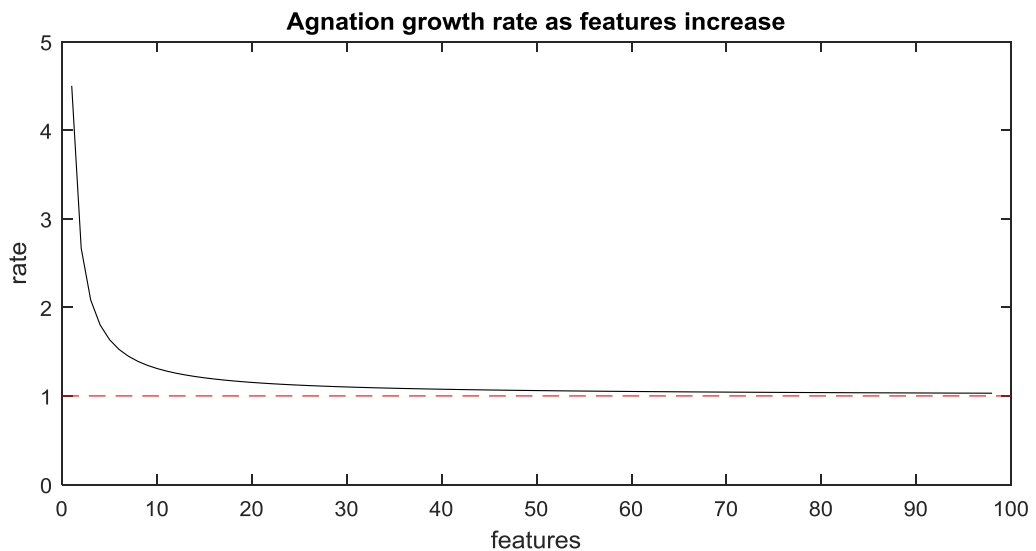


Figure 5. Agnation growth rate.

<sup>3</sup> Hypothetically, if there were a system with 1 million features, the increment growth to the previous 999,999-feature system would be 1.000000199.



## ***2.2 Quantifying delicacy and its properties***

The factors of delicacy (represented by ‘ $d$ ’) are level ( $l$ ), since delicacy is an increase in detail given by added subclassifying levels, and choice ( $e$ ), since it is possible to move from a less delicate level to a more delicate level if a choice is made. Level and choice are proportional to delicacy:  $d \propto l$ , if  $e$  is constant;  $d \propto e$ , if  $l$  is constant. Consequently, delicacy can be defined by Equation 9:

$$d = l \times e \quad (9)$$

By definition, a system can only have one level of delicacy. Any increase in detailing needs to join systems in a network. For any level  $l_i (i:1..n)$ , then  $d = \sum_{i=1}^n l_i \times e$ . Coselections of systems at the same level of delicacy are accordingly added (Figure 6).

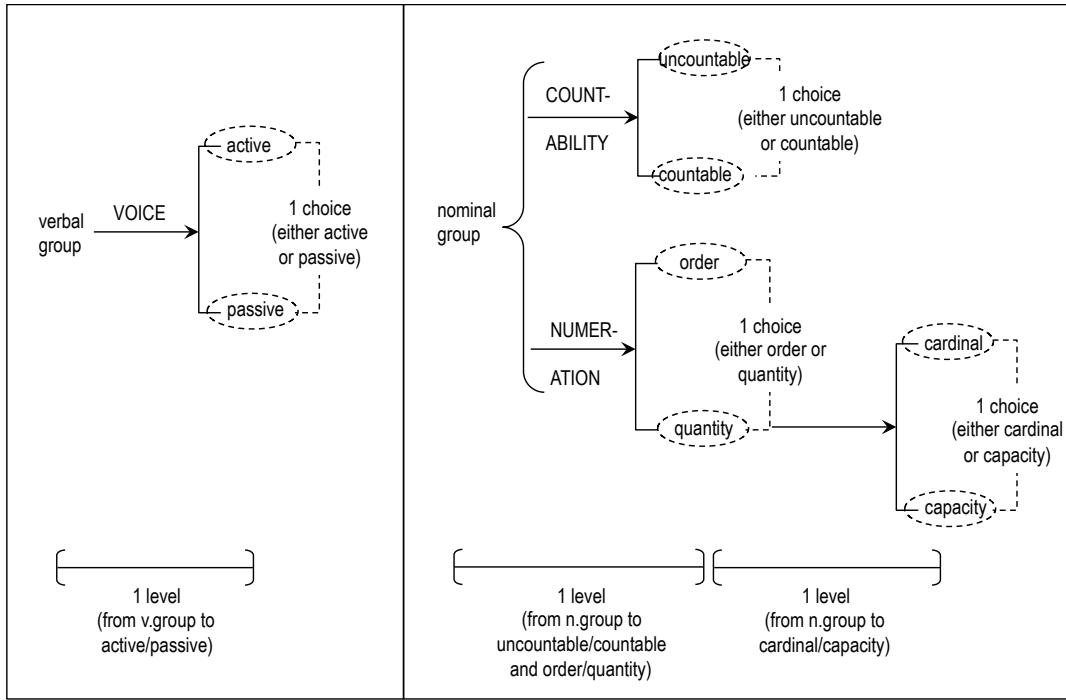


Figure 6. Comparison of delicacy between a system and a system network.

The delicacy for the system of VOICE (level = 1 and choice = 1) in English is:

$$d = l \times e = 1 \times 1 = 1 \quad (10)$$

Differently, for the network of COUNTABILITY and NUMERATION, the delicacy is define as:

$$d = \sum_{i=1}^2 l_i \times e = 1 \times 1 + 1 \times 2 = 3 \quad (11)$$

### 2.3 Quantifying valeur and its properties

Valeur (represented by 'v') is a product of agnation (a) and delicacy (d) ( $v = a \times d$ ). It is proportional to agnation and delicacy:  $v \propto a$ , if d is constant;  $v \propto d$ , if a is constant. As

such, valeur carries over the properties of its factors. From agnation, it conserves the property of asymptotic decay rate tending to 1. From delicacy, it conserves the property of summation – which means valeurs can be added up in system networks. Valeur for a network of size  $n$ , thus, can be described as:

$$v = \sum_{i=1}^n a_i \times \sum_{i=1}^n d_i \quad (12)$$

Important quantitative properties of valeur are as follows. First, simple systems have only one level of delicacy ( $n = 1$  and  $d_i$  is always equal to 1 per system), which means valeur is always equal to agnation, since  $v = a_i$ .

Secondly, valeur for system networks grows depending on the number of agnation and delicacy. For instance, if agnation remains constant (e.g., 2-feature systems keep being added to 2-feature systems), then valeur can be described as:

$$v = \sum_{i=1}^n 2 \times \sum_{i=1}^n d_i = (2 \times n) \times (n) = 2 \times n^2 \quad (13)$$

Then, the valeur of system networks also has a growth rate that decays asymptotically as the network size  $n$  increases. If we use as an example same-feature systems in a network, growing as  $a \times n^2$ , the rate of growth from  $v_i$  to  $v_{i+1}$  is as follows (Table 3):

$n$ (delicacy network size)	$2 \times n^2$ ( $a = 2$ )	$3 \times n^2$ ( $a = 3$ )	$4 \times n^2$ ( $a = 4$ )	rate ( $v_{i+1}/v_i$ )
1	2	9	24	----
2	8	36	96	4.00
3	18	81	216	2.25
4	32	144	384	1.77
5	50	225	600	1.56
6	72	324	864	1.44
7	98	441	1176	1.36
8	128	576	1536	1.30

9	162	729	1944	1.26
10	200	900	2400	1.23

Table 3. Valeur growth for  $v = a \times n^2$ .

The incremental growth in valeur is less than 2.00 for networks at delicacy ( $n = 4$ ). Taking again Caffarel *et al.* (2004) as an example, Table 4 shows the number of same-feature systems and how long their networks are:

	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$
$2n^2$	41	41	39	22	5	2	2	0
$3n^2$	1	4	0	0	0	0	0	0
$4n^2$	3	0	0	0	0	0	0	0
$5n^2$	1	0	0	0	0	0	0	0

Table 4. Delicacy for  $a \times n^2$ -system networks in Caffarel *et al.* (2004).

In real language systems, the amount of valeur tends to be constrained to networks of up to 3 or 4 systems, with a rate of 1.77 to 2.25 or more. Figure 7 is a plot of the valeur growth rate as the network increases in levels of delicacy.

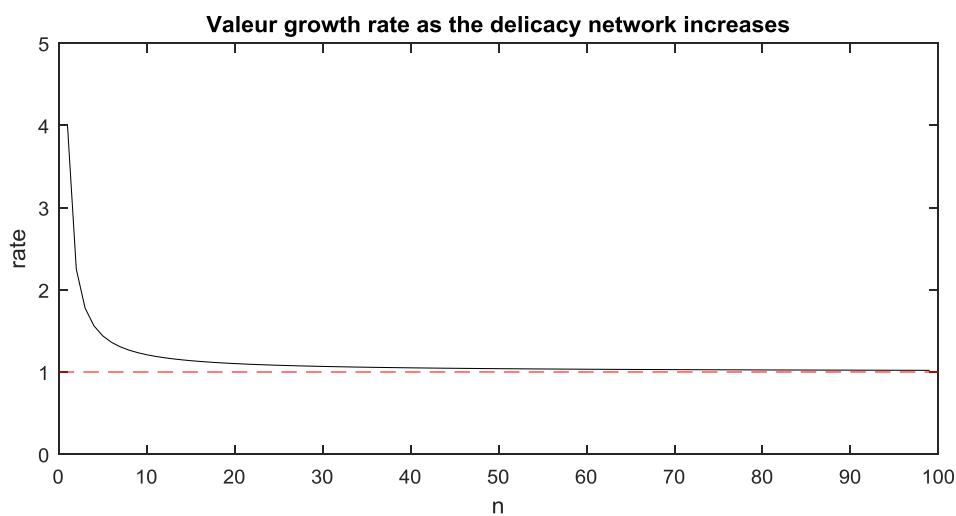


Figure 7. Valeur growth rate for  $a \times n^2$ -system networks

## *2.4 Gleason and reactance*

Since valeur is the main property of meaning, it is used as a factor to measure meaning-making precisely. The scalar total valeur, namely 'gleason' (represented by 'g') (cf. Gleason, 1965), is constructed by calculating the valeur ( $v$ ) of features in a system and calibrating them through reactances (represented by 'w') (cf. Whorf, 1945), as follows:

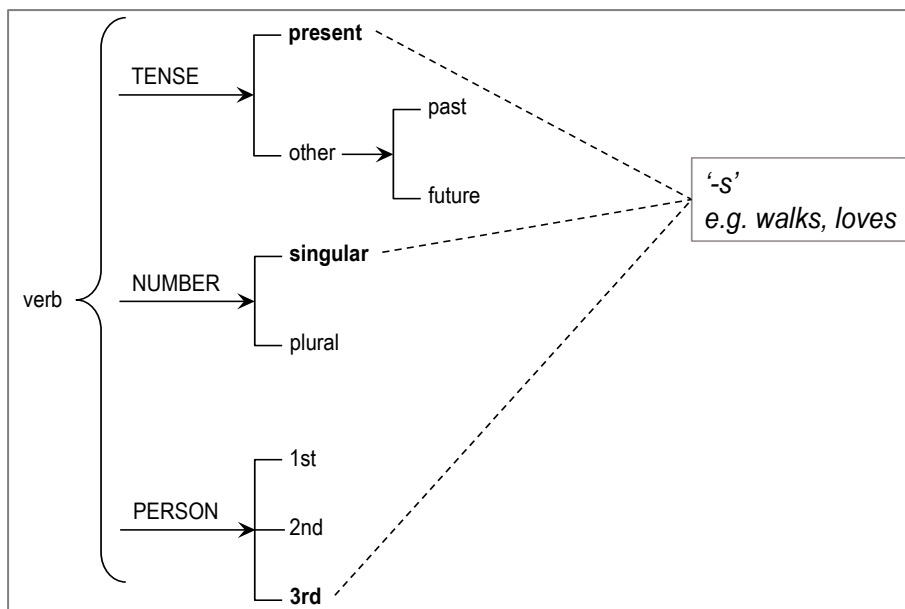
$$g = v + w \quad (14)$$

The amount of meaningfulness is therefore given by summing up valeurs and reactances. In addition, reactances ( $w$ ), as further explained below, are the bias of importance of  $v$ . Because systems co-select in actual use, valeurs and reactances are calculated considering co-selections as a system network.

Reactances are needed because features at the same level of delicacy are found to have the same valeur. They are systemically motivated and add a bias of importance to distinguish features with the same values within a network. For example, the system of TENSE in English has three features at the same level of delicacy — past, present and future. The valeur is  $v = 9$  for all three features — present ( $v = 9$ ), past ( $v = 9$ ) and future ( $v = 9$ ). When the network of systems TENSE & PERSON & NUMBER is considered, 'present' reacts differently from 'past' and 'future'. 'Present' can interact with features in PERSON and NUMBER, whereas 'past' and 'future' cannot. For example, 'present' interacts with '3rd person', but 'past' and 'future' cannot interact with '3rd person'. Accordingly, 'present' makes unique amounts of meaning when compared to the meanings 'past' and

'future' can make (Figure 8). Reactants therefore, enable the representation of those differences and the meaning potential within a context.

Figure 8. Co-selections in the verb in English.



In the network of TENSE & PERSON & NUMBER, the reactances of features are:

- For present, the reactance is equal to 2 because it can co-select with PERSON '3rd person' and NUMBER 'singular' (Figure 8).
- For features 'past' and 'future', reactances are zero (past = 0 and future = 0) because they cannot co-select.

In this context, 'present' has a richer meaning potential when compared to 'past' and 'future'. Broadly speaking, 'present' can mean more. If we compare, for instance, 'walks' and 'walked', we know that 'walks' means three things: present, 3rd person and singular (Figure 8). 'Walked', on the other hand can only mean one thing: past. 'Present', therefore, is more meaningful, as it can mean 'walk/walks' (two possibilities). Conversely, 'past' means less: 'walked' (one possibility) and 'future' also means less: 'will walk' (one possibility).

### 2.4.1 Reactance tables

The degree of interaction between systems in a network (i.e. reactances; Whorf, 1945) is precisely what is used as weight to separate out valeurs of single systems.

For each interaction (i.e., every time a feature of a system reacts to a feature from another system — e.g. present & 3rd) a weight of  $w = 1$  is summed to the valeur of that feature. So, the valeur of 'present' is 9. In the network of TENSE & PERSON & NUMBER it reacts to '3rd' and to 'singular' so a weight  $w = 2$  is assigned to 'present'. When the bias is summed to the valeur, the measure of meaningfulness of 'present' is  $g$  calculated as 9 (which is valeur  $v$ ) summed by 2 (reactances  $w$ ) which equals to 11.<sup>4</sup>

A list of reactances assigned to each feature in a network is given by comparing them to a reactance table. Reactance tables were first developed by Whorf (1945) on cryptotypes in language description. Figure 9 shows an example.

<i>location</i>			<i>commonality in features between <math>fe_i</math> and <math>fe_{i+1}</math></i>	<i>score per system</i>	
				<i>yes</i>	<i>no</i>
stratification-instantiation matrix			global semiotic address	0	1
function-rank matrix			local semiotic address	0	1
local	axis	system	entry condition	0	1
			level of delicacy	0	1
			agnations	0	1
			co-selections	0	1
			skewedness	0	1
	structure	structural function	0	1	
		realization (inter-axial)	0	1	
	rank	rank	0	1	
		class	0	1	
pre-selection		0	1		
global	stratification	context	0	1	

<sup>4</sup> It is important to note that the same process is needed to separate 'past' and 'future' out, since their valeur is 9. The same process of biasing – i.e., finding different reactances with other systems can be applied. Due to lack of space, we cannot go through the whole process here, but we can mention in passing that if we bring the system of MODALITY into the network, then 'past' reacts to modality, whereas 'future' cannot. Compare 'could/might' and '\*will can/ will may'.

		semantics	0	1
		grammar	0	1
		phonology	0	1
	instantiation	relative frequency	0	1
		register/genre generalization	0	1
		multilingual set	0	1
	metafunction	interpersonal	0	1
		ideational	0	1
		textual	0	1
		conflation	0	1

Figure 9. A table of reactances

The left-hand column 'location' places the feature in the system's architecture in terms of language dimensions — e.g. axis. The mid column details aspects of each dimension involved in the comparison of features — e.g. in axis, the place of the feature in the system as well as its realizing structure. Finally, the right-hand column scores  $fe_{i+1}$  in relation their commonality to  $fe_i$ . Each commonality is scored 'zero' and each difference is scored 'one'. The total score sum tells how apart features are. A score close to 'zero' means similar features; conversely a score distancing from 'zero' means features are more different.

### 2.5 The gleason scale

A direct consequence of using a scale is that it enables magnitudes to become units of measurement so that they can occupy a locale in space. In language, for instance, the standard unit is *g*, which measures the magnitude of “meaningfulness” of systems. For example, the meaningfulness of TENSE is given by the *gleasons* summed in the network with 'present' (11*g*), 'past' (9*g*) and 'future' (9*g*). The space occupied by an



amount of “meaningfulness” (i.e., the potential sum of *gleasons*) is named language space. Accordingly we call its spatial ordering ‘gleason scale’.

The gleason scale begins with a point (represented by ‘*fe*’, as in ‘feature’), which is a systemic feature that belongs to a given language space. The language space has 4 dimensions: ideational, interpersonal, textual (further details are given in Section 2.6) and temporal (Section 3.2). For the purpose of representation in a three-dimensional space, however, we define the feature Position (represented by ‘*P*’) as the locale occupied by a point in space relatively to other points, and is obtained through:

$$P(fe_i) = P(fe_{i-1}) + \Delta g \quad (15)$$

Where  $P(fe_i)$  is the position in the language space  $L(x, y, z)$  ( $x$  corresponding to the ideational dimension,  $y$  being the interpersonal dimension and  $z$  representing the textual dimension) of feature  $fe_i$  and  $\Delta g$  is the variation in the amount of *gleasons* between two features  $fe_i$  and  $fe_{i+1}$ .

Order (represented by ‘ord’) is the arrangement of points according to their distances from the entry condition. Consequently, any two points can be placed in order by subtracting the position of the lesser point from the greater point. Let us consider that  $fe_x$  has the least amount of *gleasons*, and  $fe_y$  is the second closest. ord therefore is calculated as:

$$ord(fe_x; fe_y) = D(P(fe_x); 0) - D(P(fe_y); 0) \Rightarrow ord(fe_x; fe_y) = fe_x - fe_y \Leftrightarrow fe_x > fe_y \quad (16)$$

Using the formalization shown above, a gleason scale can be built as a line segment in which features are points arranged in an specific order according to their

meaningfulness measured in gleasons. Figure 10 shows the system network of MENTAL PROCESS in German (Steiner & Teich, 2004) and the gleason scale for the system.

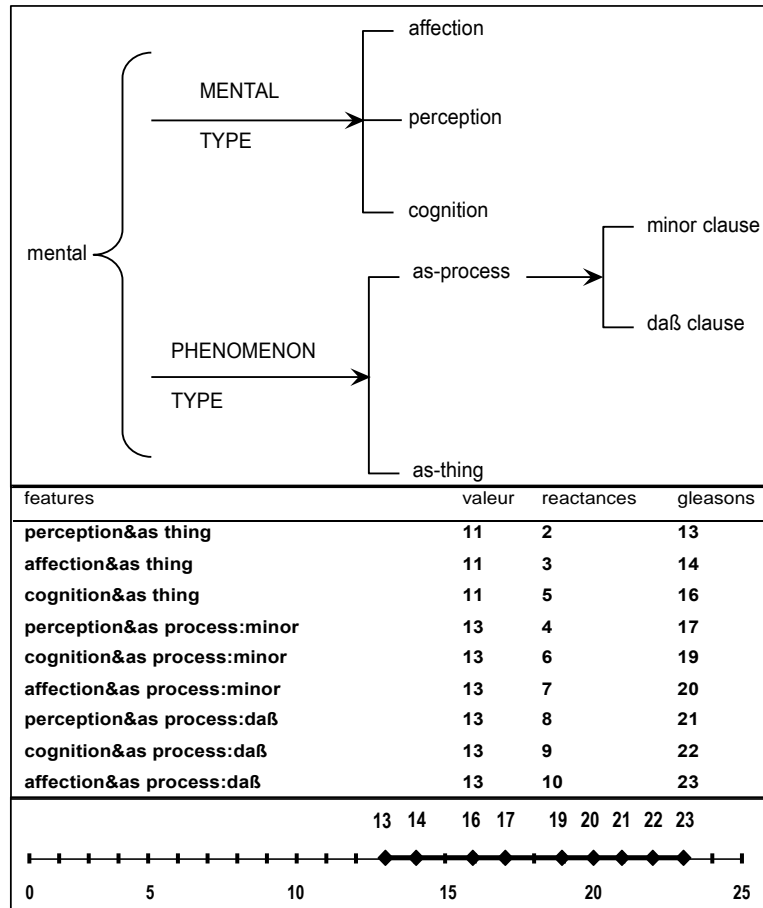


Figure 10. Gleason scale for MENTAL PROCESS in German.

The system network of MENTAL PROCESS in German has two co-selecting systems and therefore it generates nine features (middle part of Figure 10). Three of these features have a valeur of 11 and six have a valeur of 13. When reactances are considered – i.e. how this system relates to other systems in the description of German (cf. Steiner & Teich, 2004) – we can see in Figure 10 that each feature is unique and makes a unique amount of meaning, measured by the number of gleasons and ordered in the scale.

The gleason scale is fundamental to the systemic model engendered in this paper. It provides a quantitative interpretation of features in systems and scale them according to their potential to make meaning. It is through the gleason scale that we tackle the problems foregrounded in our introduction, because it affords the investigation of cryptotypical configurations quantitatively and they integrate quantitative properties to systemic functionality, since now meaning made through covert patterns and relations can be measured.

## ***2.6 Modelling meaning-making through the gleason scale***

Any language systems can be mapped onto a line segment of a gleason scale. Accordingly, a language space (represented by ‘ $L$ ’) is a model of language itself, since all language properties and behavior can be derived from the dimension of system. Consequently, all language spaces are parts of language, or ‘ $p(L)$ ’. Being a part of language, spaces may include from zero to any number of systems – which implies from zero to any number of scales. Ultimately, language can be defined as the union of all parts of language – i.e., of all language spaces with  $n \times 'p(L)'$ .<sup>5</sup>

Although there should be as many scales as there are systems, what we see in actual language use based on language descriptions is either (i) systems associating freely. Freely associating systems only go together through a contextual pressure and can be described in terms of frequencies of co-occurrence, a principle named co-variation (Martin, 2013). Or (ii) a tight bundling of systems in networks through

---

<sup>5</sup> Language spaces restricted to the grammar stratum are called grammar spaces; those restricted to discourse are called discourse spaces, and so on.

obligatory co-selections. Bundling systems cause single-system scales to collapse and form larger scales of systems.

For example, the systems of MOOD and MODALITY must co-select obligatorily as follows. MODALITY has to co-select with the feature mood: indicative, and never with mood: imperative (Figure 11). The scales of MOOD and MODALITY collapse and form a scale for the bundle.

In English, for example, “She can probably read the new book” (mood: indicative & modality: modalization), or “Read the new book!” (mood: imperative) are allowed. “Read probably the new book!” (mood: imperative & modality: modalization), however, does not exist. This is reflected on the mathematics of our model by bringing all possible gleason scales into three main scales, which then will form the language space.

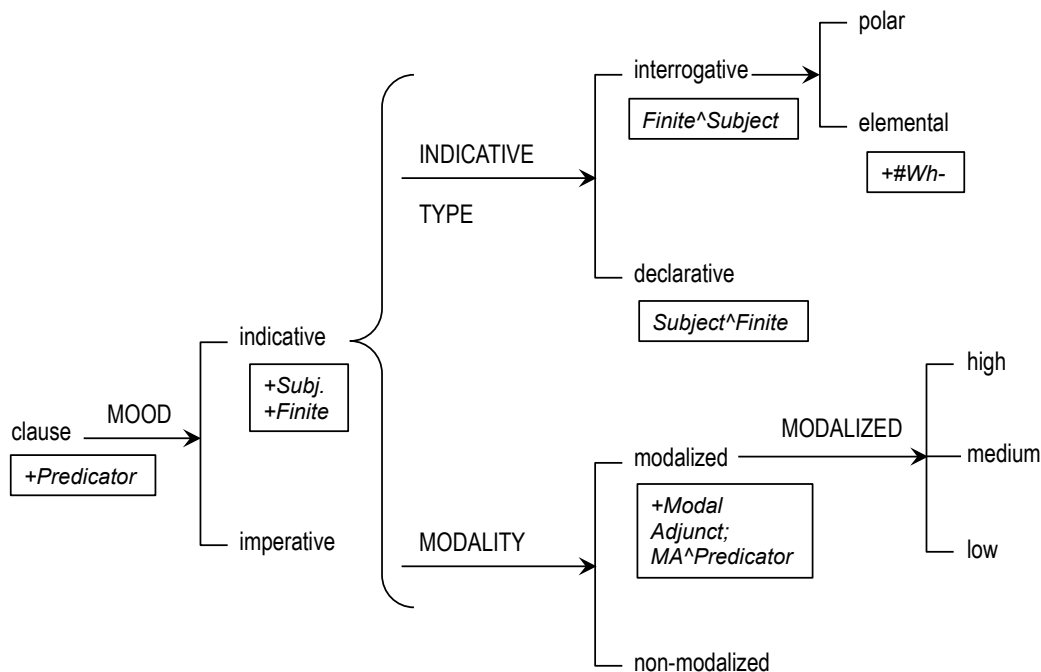


Figure 11. Co-selection of MOOD and MODALITY in English.

In SFL, system bundling is named metafunction. In every language described so far, three metafunctions have been identified (Halliday 1970; Halliday & Matthiessen, 2014). Interpersonal metafunction is responsible for making meanings related to enacting social relations; ideational metafunction is responsible for making meanings related to representation of events and things; textual metafunction is responsible for making meanings related to composing text. Figure 12 lists main systems bundled by metafunction commonly found in different languages.

metafunction	systems bundling from different strata
interpersonal	MOOD, MODALITY, ASSESSMENT, POLARITY, NEGOTIATION, COMMENT, APPRAISAL, SPEECH FUNCTION, EXCHANGE, TONE...
ideational	TRANSITIVITY, CIRCUMSTANCE, TAXIS, LOGICO-SEMANTICS, EXTERNAL CONJUNCTION, RHETORICAL RELATIONS...
Textual	THEME, PREDICATION, COHESION, INTERNAL CONJUNCTION, MESSAGE, INFORMATION, TONICITY...

Figure 12. Systems bundled by metafunctions.

As a result, modelling meaning-making ends up with a gleason scale per metafunction, which can then co-variate (i.e., associate freely) depending on contextual pressure and use. Metafunction-based scales form the language space, which houses all possible features for any systems mapped onto them. Accordingly, the language space allows all possible co-variations. Figure 13 shows a visualization of the language space with scales of ten features. This space holds a thousand co-variations in meaning-making.

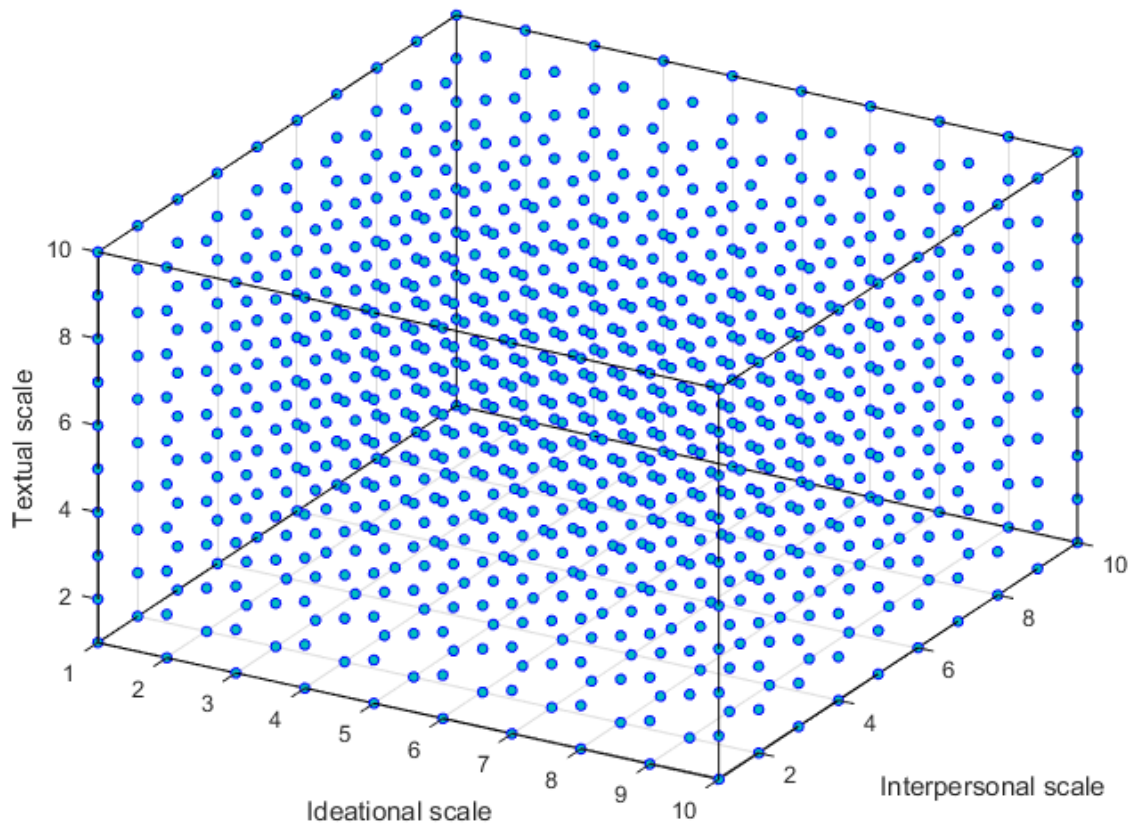


Figure 13. Language space of 10 x 10 x 10.

Figure 13 is just an illustration. Actual language spaces constructed from language descriptions have many more possibilities. If we take the grammar of English as an example (Halliday & Matthiessen, 2014), and restrict Gleason scales to major clause systems (TRANSITIVITY & CIRCUMSTANTIATION, MOOD & MODALITY and THEME & PREDICATION), we get a space of 30,011 x 221,532 x 1136 – which holds 7,552,578,823,872 (over  $7.5 \times 10^{12}$ ) possible co-variations in meaning making<sup>6</sup>.

---

<sup>6</sup> It is worth noting that obviously not all co-variations occur in actual language use; in fact, the vast majority don't and probably never will. However, the language space – being the product of evolution – shows how enormous the potential of language to make meaning is.

### 3. Modelling text unfolding

The methods section has shown how to find valeur in systems, then order systemic features in scales and build language spaces. Now we are in a position to describe the unfolding of meaning-making in text.

As an example, we will model a text extracted from a corpus called Popular Science Modelling, composed by texts which objective is to explain science facts to children. As an example for the present modelling, we will use a text called ‘Balloon speakers’, which teaches children to do a simple science experiment on sound amplification<sup>7</sup>.

#### 3.1 Text profiling

The modelling of meaning-making in the language space involves (1) static text profiling and (2) dynamic text unfolding.

Static text profiles matches features to coordinates in the language space. They provide a description of language systems used in a text measured in *gleasons*. Text profiles may involve scales of single systems up to all metafunctional spaces.

For the text profiling of ‘Balloon speakers’ clause grammar systems were selected. TRANSITIVITY & CIRCUMSTANTIATION for ideational metafunction; MOOD & MODALITY for interpersonal metafunction and THEME & PREDICATION for textual metafunction. The text was divided into units of analysis – for the grammar, ‘clause’

---

<sup>7</sup> Collected from [www.sciencekids.co.nz](http://www.sciencekids.co.nz).

was chosen. Systemic features were identified and mapped onto the gleason scales using the calculations developed above in Section 2.

Using clauses 1 and 2 as an example, the gleasons were calculated shown in Table 5.

Table 5. Grammatical analysis of 'clauses 1 and 2' for the clause systems.

<b>clause</b>	<i>system</i>	<i>features</i>	<i>gleasons</i>
<b>1 Use a balloon</b>	TRANSITIVITY	material: transformative: extending & transitive: operative	232
	CIRCUMST.	non-selected	0
	MOOD	clause: major: free: positive & minus assessment & interactant: addressee & imperative: implicit	25
	MODALITY	non-selected	0
	THEME	predicator theme: imperative: identified & minus interpersonal & minus textual	9
	PREDICATION	non-selected	0
<b>2 to amplify sound</b>	TRANSITIVITY	material: transformative: extending & transitive: operative	232
	CIRCUMST.	non-selected	0
	MOOD	clause: major: bound: positive & minus assessment	9
	MODALITY	non-selected	0
	THEME	non-selected	0
	PREDICATION	non-selected	0

Table 6 shows the points in the gleason scales for the features identified in each clause.

Table 6. Text profiling of 'Balloon Speakers'

<b>Clause</b>	<b>Balloon speakers</b>	<b>ideational</b>	<b>interpersonal</b>	<b>Textual</b>
1	Use a Balloon	232	25	9
2	to Amplify Sound	232	9	0



3	Small sounds can still make a big noise	232	297	65
4	when you use a good sound conductor.	232	17	67
5	Experiment with a balloon, compressed air and your own ears	232	25	9
6	to find out	232	9	0
7	how it works and the science behind it.	209	297	66
8	What you'll need:	232	1449	125
9	Balloon	0	9	0
10	Instructions:	0	9	0
11	Blow up the balloon.	232	25	9
12	Hold the balloon close to your ear	232	25	9
13	while you tap lightly on the other side.	209	17	66
14	What's happening?	232	729	87
15	Despite you only tapping lightly on the balloon	232	9	66
16	your ears can hear the noise loudly.	1265	297	65
17	When you blew up the balloon	232	17	66
18	you forced the air molecules inside the balloon closer to each other.	830	873	65
19	Because the air molecules inside the balloon are closer together,	2529	17	66
20	they become a better conductor of sound waves than the ordinary air around you.	2483	297	65

Finally, each point is represented within the grammar space. This gives a static profiling of the text. Starting from the origin 0 ideational (x-axis), 0 interpersonal (y-axis) and 0 textual (z-axis), we can measure precisely how far each point is from the origin in each axis.

Figure 14 shows that some clauses appear very close to each other (e.g. clauses 7 and 3), or even overlap (e.g. clauses 9 and 10). Accordingly, some points tend to be less distant to one another, while others tend to be more distant. This is explained by the number of features each clause is making per clause and the resources language needs to make them. Distance between points will be explained in relation to text dynamics in Section 4 below.

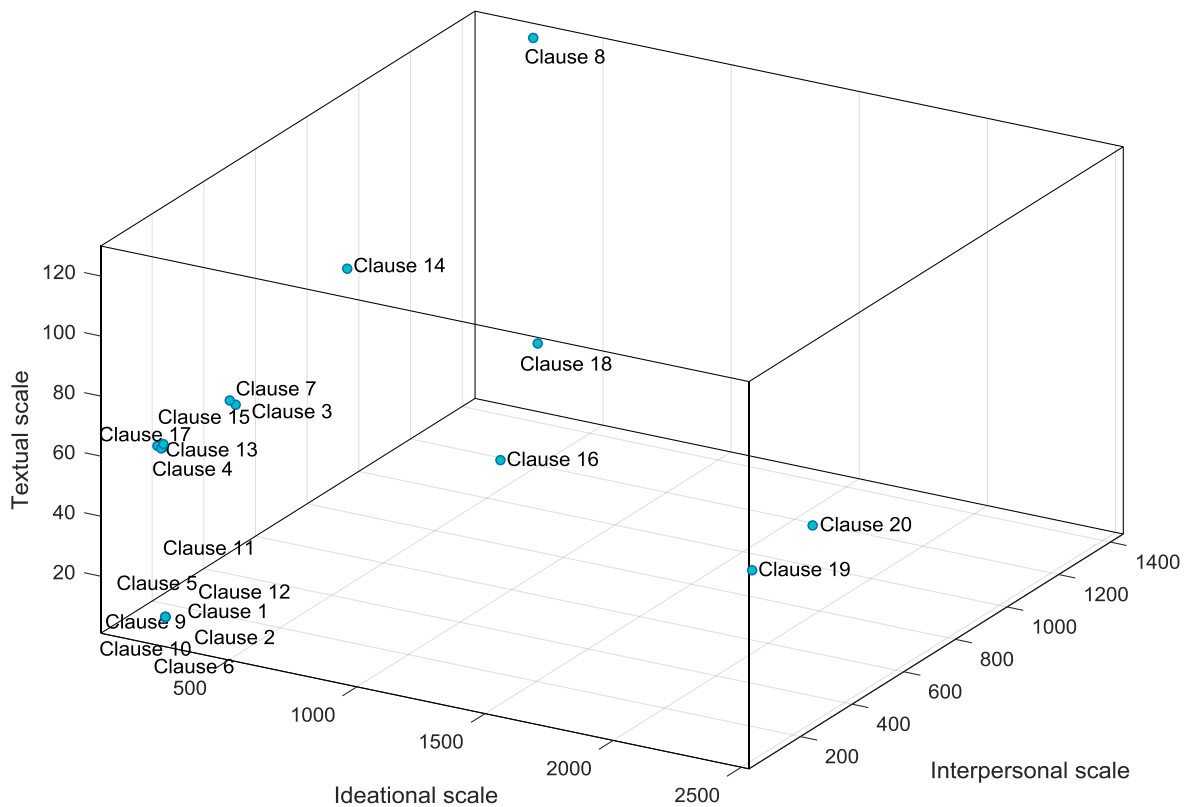


Figure 14. ‘Balloon speakers’ in the grammar space of English.

### 3.2 Text dynamics

Dynamics involves two important concepts – ‘time’ and ‘movement’. A text is the final product of a choice made on the system network realized by some structure which is then followed by another choice and so on. Every choice in system causes a displacement in the language space, a movement from one point to another. ‘Choice-in-network’ can be used as a measure of time – or movement of ‘text in the making’ in the

language space. We can therefore devise a ‘semiotic particle’, as it were, to trace text displacement in space.

Introducing time leads to understanding text both as (i) the accumulation of meaning as the ‘semiotic particle’ moves along – interpreted globally as the text as a whole; and (ii) the movement from time ‘ $t_i$ ’ to time ‘ $t_{i+1}$ ’ – interpreted locally as movement.

The (i) accumulation of meaning is given by the fact that every system network has a language unit as entry condition, such as phonemes, syllables, words, figures and negotiations. In the grammar space, the entry condition for the systems analyzed in this paper is the clause. Clause therefore is our unit of time. The meanings of a text are the accumulating result of the meanings at each clause.

Figure 15 shows the amount of meanings made in ‘Balloon speakers’ per metafunctional scale and globally, measured in *gleasons*. There are 20 clauses in the text, as shown previously in Table 5. The text made therefore 10,309g in the ideational scale, 4,452g in the interpersonal scale, 903g in the textual scale and 15,666g globally (sum of each scale for the clauses in Table 5).

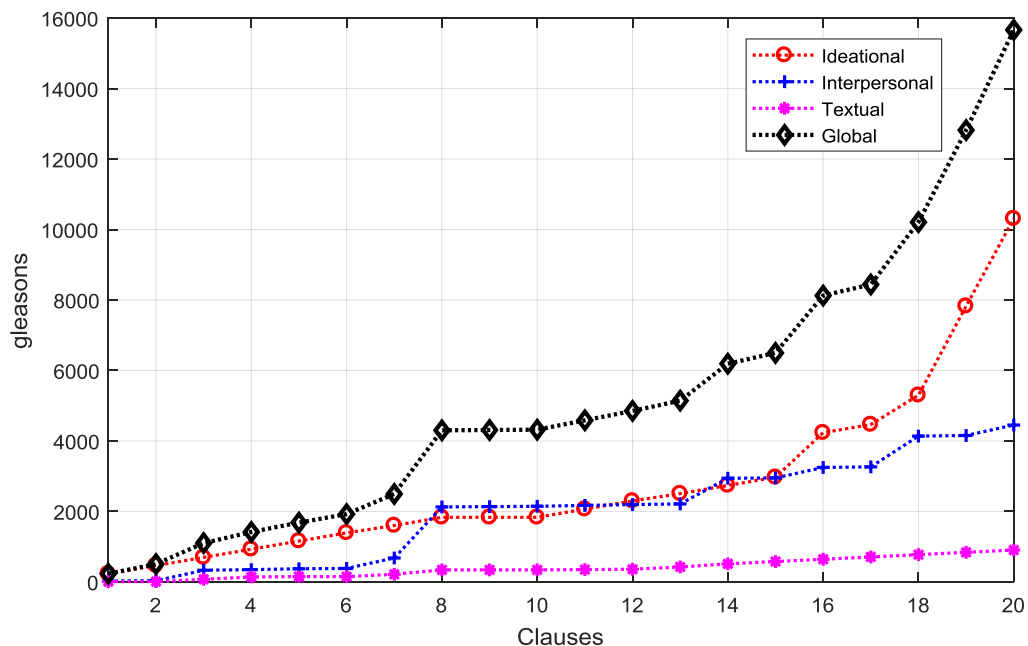


Figure 15: Amount of *gleasons* in ‘Balloon Speakers’ as the text evolves in time for ideational, interpersonal, textual and global scales.

The (ii) local movement is given by the fact that for every unit of time (i.e., for every clause) we can measure the distance between two sets of coordinates and calculate how many meanings were made from a clause to the next. By modelling the movement of language in space, we can use the *gleason* scales as parameters to track displacement.

‘Balloon speakers’ begins at the origin, where no meaning is made. From there to the first clause ( $t = 1$ ), 232g were made in the ideational scale (x-axis); 25g in the interpersonal scale (y-axis) and 9g in the textual scale (z-axis) (see Table 5, second line regarding clause 1). The second clause of the text has  $x = 232g$ ,  $y = 9g$ ,  $z = 0g$ .

Since the *gleason* is a scalar and semiotic movement is displacement in space, it can be represented by the line segment between two points in the gleason scales of each axis. Using parametric equations for each coordinate in the language space, we can represent movement ( $r$ ) at each instance of time. The whole account is presented in Table 7.

Table 7: Movement in ‘Balloon Speakers’.

<b>movement</b>	<b>r = clause (x, y, z)</b>		<b>rx = idea. (x<sub>1</sub> - x<sub>0</sub>)</b>	<b>ry = interp. (y<sub>1</sub> - y<sub>0</sub>)</b>	<b>rz = text. (z<sub>1</sub> - z<sub>0</sub>)</b>
r <sub>1</sub>	from 0	to 1	232	25	9
r <sub>2</sub>	from 1	to 2	0	-16	-9
r <sub>3</sub>	from 2	to 3	0	288	65
r <sub>4</sub>	from 3	to 4	0	-280	2
r <sub>5</sub>	from 4	to 5	0	8	-58
r <sub>6</sub>	from 5	to 6	0	-16	-9
r <sub>7</sub>	from 6	to 7	-23	288	66
r <sub>8</sub>	from 7	to 8	23	1152	59

r9	from 8	to 9	-232	-1440	-125
r10	from 9	to 10	0	0	0
r11	from 10	to 11	232	16	9
r12	from 11	to 12	0	0	0
r13	from 12	to 13	-23	-8	57
r14	from 13	to 14	23	712	21
r15	from 14	to 15	0	-720	-21
r16	from 15	to 16	1033	288	-1
r17	from 16	to 17	-1033	-280	1
r18	from 17	to 18	598	856	-1
r19	from 18	to 19	1699	-856	1
r20	from 19	to 20	-46	280	-1

---

If we consider all axes together and displacement in the language space, we see the movement between points in space. The difference in *gleasons* between meanings made at each unit of time show dynamically the amount of meanings made throughout the text (Figure 16).

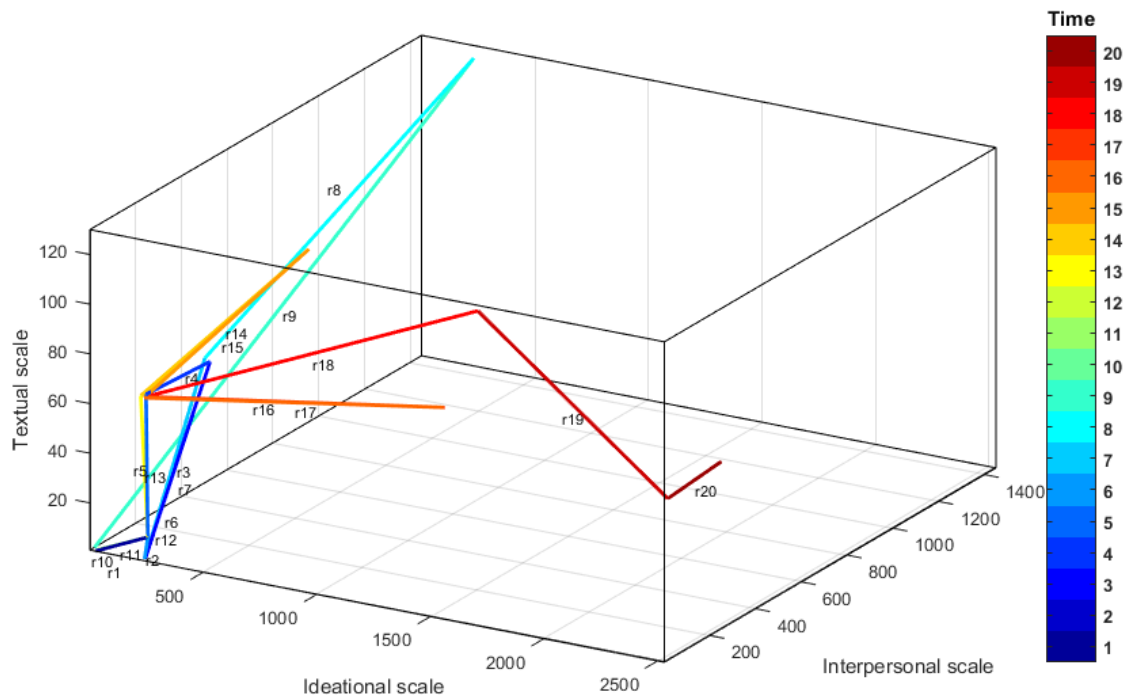


Figure 16. A systemic dynamics model of the meaning-making process of ‘Balloon Speakers’.

#### 4. Discussion

The model presented in this paper is a representation of meaning-making in the language space involving a static text profiling and a dynamic text unfolding.

Static text profiling in ‘Balloon speakers’ showed how clauses disperse in the grammar space. It used the gleasons of each scale as coordinates for points representing the dispersion. The distance between points is a measure of similarity among resources being deployed. Points which are closer to one another indicate similar resources, while distant points indicate that resources are different.

The grammar systems in ‘Balloon speakers’ tend to deploy similar resources to realize similar textual functions – such as material actions for procedures or non-

modalized declaratives for explanations. From the point of view of text, this is the reason why clauses group together as sequences.

Accordingly, this is also the reason why points of similar coordinates tend to form groupings in space. 'Balloon speakers' is popular science teaching kids how to make simple science experiments, then explaining the science behind it. The results show that the two groupings in the text are an experiment-grouping and an explanation-grouping.

The experimental-grouping (points 1, 2, 5, 6, 9, 10, 11, 12, 13) is material actions and imperative mood, non-modalized clauses, e.g. "use a balloon" and "blow up the balloon". The explanation-grouping (points 4, 13, 15, 17, 19) is material actions, but dependent in mood. These give some extra explanation about the steps of the experiment, e.g. "when you use a good sound conductor" and "while you tap lightly on the other side".

The points which do not group together perform specific jobs in the text such as change phase from introduction to experiment (clause 8); interact with the reader (clause 14) or relate concepts in science explanation (clause 20).

Dynamic text unfolding in 'Balloon speakers' therefore showed both a global accumulation of meaning and the local movement at each moment in time.

Global accumulation shows how systems/metafunctions are contributing to the overall meaning in 'Balloon speakers'. The ideational metafunction contributes more in the beginning of the text, from  $t = 1$  to  $t = 7$ . Then, from  $t = 8$  to  $t = 15$ , it is the interpersonal metafunction that contributes more (Figure 15). Finally, from  $t = 16$  on, the ideational metafunction contributes more.

The amount of meanings made in a text leads to important questions about text complexity. It can, for instance, correlate why some texts "feel" easier or harder to

variation in gleasons. This may help building a text typology based on the amount of gleasons of each text type and their pattern of variation per type. Dynamically, the amount of meaning made in a given period of time can be used as criterion to classify texts as similar or different. It may also correlate text complexity and human activity complexity, where more complex human activities demand more complex texts, which in turn demand more gleason variation in meaning-making.

Local movement shows how many extra meanings need to be made or lost at each point in text. Knowing the amount of gleasons made for combinations of sequence of clauses can be a tool to understand the most probable combinations of points. In this sense, dynamic analysis can be predictive of what is coming next in a text.

Predictability of the potential meanings a text can make is potentially an important tool for many applications, particularly in contexts of language education by designing a curriculum based on predictions about patterns of text unfolding, or in translation text machine learning.

More importantly, dynamics can show the direction the ‘semiotic particle’ is taking in the language space and the movement afforded at any moment in time. From this point on, a text is not seen as a structure anymore. Rather, it is seen as movement; a variation in the meaning making-process between units of time. So, instead of labelling a meaning made by a system as, for instance, a material:action & declarative:non-modalized & theme:simple, it can be modelled as the movement in space as  $r_4 = 1 (0_4, -280_4, 2_4)$ .

## **5. Conclusion**



As results produced in Quantitative Linguistics equip the field to expand the scope of research, quantitative methods can be applied to problems of growing complexity. We addressed one of these problems here; namely, to develop quantitative methods to explain language organization as dynamic patterns of choice and how these are deployed in actual language use.

Because actual language use can only be observed in texts dynamically, the model conceptualized text as the resulting process of systemic organization. Consequently, 'time' as variation in choice at each moment in text was introduced.

The model drew from SFL (Halliday & Kress, 1976) to identify text's systemic properties and model them quantitatively. Systems were modelled in terms of their essential quantitative property – valeur.

We then proceeded to calculate valeur and add it to reactances to form gleasons. Gleasons, in turn, were defined as a semiotic unit and placed in a scale, enabling features to be ordered. Because systems co-select, gleason scales can become intersecting axes, thus forming a space. Having determined a grammar space, we tracked the meaning-making process using the text 'Balloon speakers' as an example of application. We were able to determine how meanings were being made dynamically and to produce a model.

Position of points in space revealed the function of groups of clauses in the text. The text becomes a particle moving in the language space, enabling the interpretation of points in space as a form of text analysis – and more importantly one that can be achieved by exact methods.

The model has proven to be useful to model the example text. It may, consequently, inform future research. It showed that dynamic accumulation is useful to identify core characteristics of a text by revealing how each group of systems is mostly

contributing to make meaning. This may help future research investigating prediction of movement in text.

Finally, the model shows how local movement gives a precise variation in meaning for a particular moment in text. Assuming that similar texts obey similar patterns of movement, dynamic modelling may be further explored in future research on text typology and text recognition, which may be used in applications in education, translation and machine learning.

## 6. References

- Altmann, G. (2015) *Studies in Quantitative Linguistics* 21. Lüdenscheid: RAM-Verlag.
- Caffarel, A., Martin, J., & Matthiessen, C. (Eds.). (2004). *Language Typology: A functional perspective*. Amsterdam: John Benjamins Publishing Company.
- Dione, C. (2014). LFG parse disambiguation for Wolof. *Journal of Language Modelling*, 2(1), 105–165.
- Gleason, H. (1965). *Linguistics and English Grammar*. New York: Holt, Rinehart & Winston Ltd.
- Halliday, M. Language structure and language function. In: LYONS, J. (Ed.). *New Horizons in Linguistics*. Harmondsworth: Penguin Books, 1970.
- Halliday, M. (1978). *Language as social semiotic: the social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.; Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M.; James, Z. (1993). A quantitative study of polarity and primary tense in the English finite clause. In: Sinclair, J.; Hoey, M.; Fox, G. (Eds.). *Techniques of description: spoken and written discourse*. London and New York: Routledge. 32–66.

- Halliday, M., & Kress, G. (1976). *Halliday: system and function in language: selected papers*. London: Oxford University Press.
- Halliday, M.; Matthiessen, C. (4th Edition) (2014). *An Introduction to Functional Grammar*. London & New York: Routledge.
- Halliday, M.; Webster, J. (2005). *Computational and Quantitative Studies*. London & New York: Continuum.
- Ke, S. (2012). Clustering a Translational Corpus. In Oakes, M., Ji, M. (Eds.). *Quantitative Methods in Corpus-Based Translation Studies: A Practical Guide to Descriptive Translation Research* (p. 149-174). Amsterdam: John Benjamins.
- Lemke, J. (1984). *Semiotics and education*. Toronto: Toronto Semiotic Circle.
- Lemke, J. (1991). Text Production and Dynamic Text Semantics. In Ventola, E. (Ed). *Functional and Systemic Linguistics: Approaches and Uses*. p. 23-38. Berlin: Mouton/de Gruyter.
- Lemke, J. (1993). Discourse, Dynamics, and Social Change. *Cultural Dynamics*, 6(1), 243-275.
- Martin, J. (1992). *English Text: System and structure*. Amsterdam: John Benjamins Publishing Company.
- Martin, J. (2013). *Systemic functional grammar: a next step into the theory – axial relations*. Beijing: Higher Education Press.
- Matthiessen, C. (1995). *Lexicogrammatical cartography: English systems*. Tokyo: International Language Sciences Publishers.
- Matthiessen, C. (1999). The system of TRANSITIVITY: an exploratory study of text based profiles. *Functions of Language*, 6(1), 1–51.

- Naumann, S. (2015). Syntactic Complexity in Quantitative Linguistics. In Tuzzi, A., Benešová, M.; Macutek, J. (2015). *Recent Contributions to Quantitative Linguistics*. Berlin, Boston: De Gruyter Mouton.
- Nesbitt, C. and Plum, G. (1988) 'Probabilities in a systemic functional grammar: the clause complex in English,' Fawcett, R., Young, D. (eds). *New Developments in Systemic Linguistics, Volume 2: Theory and Applications*. London: Frances Pinter.
- Plum, G. & Cowling, A. (1987). Social constraints on grammatical variables: tense choice in English', in Steele, R., Threadgold, T. (Eds). *Language Topics*. Vol II. Amsterdam / Philadelphia: John Benjamins, 281-305.
- Popescu, I.-I.; Lupea, M.; Tatar, D.; Altmann, G. (2015). *Quantitative Analysis of Poetic Texts*. Berlin, Boston: De Gruyter Mouton.
- Popescu, I.-I., Mačutek, J., Altmann, G. (2009). *Aspects of Word Frequencies*. Lüdenscheid: RAM-Verlag.
- Rybicki, J., Eder, M. (2011). Deeper Delta across genres and languages: do we really need the most frequent words? *Literary and Linguistic Computing*, 26(3), 315–321.
- Rybicki, J., Heydel, M. (2013). The stylistics and stylometry of collaborative translation: Woolf's Night and Day in Polish. *Literary and Linguistic Computing*, 28(4), 708–717.
- Saussure, F. (1966). *Course in general linguistics*. New York: McGraw-Hill Book Co.
- Steiner, E., & Teich, E. (2004). Metafunctional profile of German. In Caffarel, A., Martin, J., Matthiessen, C. (Eds.). *Language Typology: A functional perspective* (Vol. 253). Amsterdam: John Benjamins Publishing Company.
- Tuzzi, A., Benešová, M., & Macutek, J. (2015). *Recent Contributions to Quantitative Linguistics*. Berlin, Boston: De Gruyter Mouton.
- Whorf, B. (1956). *Language, thought, and reality: selected writings*. Cambridge, Mass.: The MIT Press.

Zhang, H., Liu, H. (2017). Motifs in reconstructed RST discourse trees. *Journal of Quantitative Linguistics*, 24.

## 7. List of Tables

Table 1. Feature per system in Caffarel *et al.* (2004).

Table 2. Agnation growth.

Table 3. Valeur growth for  $v = a \times n^2$ .

Table 4. Delicacy for ad<sup>2</sup>-system networks in Caffarel *et al.* (2004).

Table 5. Grammatical analysis of 'clauses 1 and 2' for the clause systems.

Table 6. Text profiling of 'Balloon Speakers'

Table 7: Movement in 'Balloon Speakers'.

## 8. List of Figures

Figure 1. A system.

Figure 2. The system of VOICE in English.

Figure 3. Cryptotypical meaning of "mass noun" in English.

Figure 4. Agnation and delicacy.

Figure 5. Agnation growth.

Figure 6. Comparison of delicacy between a system and a system network.

Figure 7. Valeur growth for ad<sup>2</sup>-system networks

Figure 8. Co-selections in the verb in English.

Figure 9. A table of reactances

Figure 10. Gleason scale for MENTAL PROCESS in German.

Figure 11. Co-selection of MOOD and MODALITY in English.

Figure 12. Systems bundled by metafunctions.

Figure 13. Language space of 10 x 10 x 10.

Figure 14. The text 'Balloon speakers' in the grammar space of English.

Figure 15: Amount of gleasons in 'Balloon Speakers' – ideational, interpersonal, textual and global.

Figure 16. A systemic dynamics model of the meaning-making process of 'Balloon Speakers'.