Cohesive Harmony Analysis: Measuring the Coherence of a Text

TIAN Jianqiu. Peking University, China

Absrtact: Hasan's (1984) Model of Cohesive Harmony Analysis (CHA) represents one of the first attempts within the SFL framework to devise a relatively objective index for measuring a text's coherence. As a development of the classic "cohesion analysis" (Halliday & Hasan 1976), it probes into the coherence of a text by considering discourse structure as being based in both the syntagmatic and the paradigmatic semantic relations between components of clauses. More than thirty years later, it is still widely used as an effective tool for language teaching and linguistic pathology. This paper will trace its origin, introduce its basic concepts and analytical procedure, describe its applications and consider its possible future development.

1. Origin

H asan's Model of Cohesive Harmony Analysis was a continuation of research on cohesion analysis.

Cohesion analysis originates from Hasan's work on the Nuffield Programme in Linguistics and English Teaching at University College London (1965-1967). The research identified the majority of linguistic resources for creating continuity in English texts, which formed a large part of the classic Cohesion in English (Halliday & Hasan 1976). The linguistic resources have been called cohesive devices.

Later research that gave rise to Cohesive Harmony Analysis was conducted at the Sociological Research Unit at the University of London Institute of Education under the general directorship of Basil Bernstein (Hasan 1984). One of the aims of the research project was to determine how close children of varying social backgrounds came to telling a bedtime story when asked to do so. To be specific, they were to check if there existed any correlation between certain social factors and the degree of coherence perceived in the extempore texts, produced by 6 to 7 years old

About the Author: TIAN Jianqiu,associate professor in the School of Foreign Languages at Peking University, China. She received her MA in linguistics 1997 and is currently working on her PhD in linguistics from the same university. Her research interests include genre study, context in Systemic Functional Linguistics and discourse analysis.

children from different social backgrounds. Typical examples were as follows:

1)I have a cat. It loves liver. It sleeps all day. It is black.

2)A cat is sitting on a fence. A fence is made of wood. Carpenters work with wood. Wood planks can be bought from a timberyard.

3)The captain has made a mistake. He will marry the female and bury her in an empty hole. He felt faint. So he sat against the drain which was under repair. The Enemy were trapped. So the taxis had to hurry to the pleasant grassy slopes to save them.

From an examination of the texts, Hasan pointed out that whatever method of analysis we adopt, 3) stands out as the longest piece and yet the least coherent of the three. From these observations, Hasan made the following assumptions of coherence:

i. that normal speakers are sensitive to variation in coherence;

ii. that textual coherence is a relative, not an absolute property, so that it is possible to rank a group of texts on a cline (Halliday 1961) from most coherent to least coherent; and as the membership of the group of texts changes, so might the position of individual texts on the cline;

iii. that coherence is an essential property of texts; consider that normal speakers of English would not regard 3) as a text;

iv. that wherever a textual fragment exceeds one simple sentence the variation in coherence does not correlate with structural facts; such structure is only a necessary but insufficient condition;

v. that, therefore, an examination of coherence necessarily involves an examination of nonstructural relations of the type that we refer to as cohesion.

With these assumptions in mind, this research examined important parameters for the classification of cohesive chains and their functions in the ecology of a text (Hasan 1973). Hasan suggested that cohesive ties are not sufficient conditions for the coherence of a text; coherence requires a particular type of dual relation between elements of two or more cohesive chains: on the one hand these items were cohesive by virtue of being in a cohesive chain, on the other their cohesive power was enhanced by an experiential relation based largely but not exclusively on transitivity functions. This dual relation formed the basis of what Hasan called cohesive harmony.

2. Basic concepts

The following concepts are crucial for discussing cohesive harmony, texture and coherence.

2.1 Cohesive tie

In talking about texture, the concept that is most important is that of a TIE. The term itself implies a relation: you cannot have a tie without two members, and the members cannot appear in a tie unless there is a relation between them. Let us draw a picture of the tie:

A   B
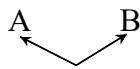
If you think of a text as a continuous space in which individual messages follow each other, then the items that function as the two ends of the tie - the A and the B - are spatially separated from each other; A may be part of one message and B part of another. But there is a link between the two, depicted above by the two-headed arrow. The nature of this link is semantic: the two terms of any tie are tied together through some meaning relation. Such semantic relations form the basis for cohesion between the messages of a text. There are certain kinds of meaning relation that may obtain between the two members. What ties the two members of a tie are basically three semantic relations: co referentiality, co-classification, and co-extension. The existence of such ties is essential to texture; the longer the text, the truer this statement.

2.2 Cohesive device

The linguistic resources that encode the semantic relations making cohesive ties are called cohesive devices, which Hasan (1985) summarized in Figure 1.

2.3 Cohesive chain

In a typical text, grammatical and lexical cohesion move hand in hand, the one supporting the other. Many different kinds of semantic relations operate at one and the same time through sizeable portions of a text. To demonstrate this point, let me examine in some detail the first five clauses of Text 1:

    Text 1:
    1) once upon a time there was a little girl
    2) and she went out for a walk

COMPONENTIAL RELATIONS		ORGANIC RELATIONS
Device	Typical tie relation	**A: Conjunctives** e.g. causal tie concession tie ...

Reference
1. Pronominals
2. Demonstratives
3. Definite article
4. Comparatives
	co-reference	**B: Adjacency pairs**
		e.g. Question (followed by)
		    answer;
		    offer (followed by)
		    acceptance;
		    order (followed by)
		    compliance ...

Substitution & **Ellipsis** 1. Nominal 2. Verbal 3. Clausal	co-classification	
General 1. Repetition 2. Synonymy 3. Antonymy 4. Meronymy	co-classification or co-extension	**Continuatives** (e.g. still, already...)
Instantial 1. Equivalence 2. Naming 3. Semblance	co-reference or co-classification	

STRUCTURAL   COHESION

Parallelism
Theme-Rheme Development
Given-New Organisation

Figure 1 Summary of cohesive devices

3) and she saw a lovely little Teddybear

4) and so she took it home

5) and when she got home she washed it

In Figure 2 each rectangle stands for one-clause. Within each of these clauses there are components that enter into a grammatical or lexical cohesive relation. There are four such threads of continuity:

1) the first, with the first element girl in clause 1;

2) the second, with went in clause 2;

3) the third, with teddy bear in clause 3; and
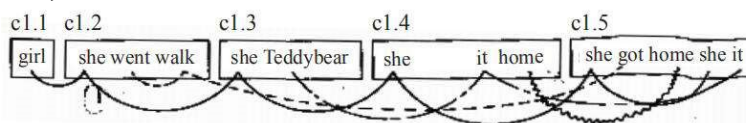
4) the fourth, with home in clause 4.

Figure 2 cohesive chains in Text 1

As the analysis provided in Figure 2 shows, a cohesive chain is formed by a set of items each of which is related to the others by the semantic relation of co-reference, co-classification, and/or co-extension.

Taking the type of relation into account, we can sub-categorise chains into two types: IDENTITY CHAINS and SIMILARITY CHAINS. Again, both of these are exemplified in Figure 2. Thus chain 1 with girl, she, etc. is an identity chain. The relation between the members of an identity chain is that of co-reference: every member of the chain refers to the same thing, event, or whatever, as in this chain, where each item refers to the same girl.

An example of similarity chains is provided by chain 2 in Figure 2 with went, walk, etc.: the members of a similarity chain are related to each other either by co-classification or co-extension. Each such chain is made up of the same class of things, events, etc, or to members of non-identical but related classes of things, event, etc.

2.4 Chain interaction

Chain interaction means relations that bring together members of two (or more) distinct chains. These relations are essentially grammatical. For example, if we take chain (a) girl and chain (e) went, walk, got from Figure 3, we would note that girl is in an identical grammatical relation with went and got: girl is the ACTOR of the ACTION went and got.

Identity chains
- (a) girl (17)
- (b) Teddybear (14)
- (c) home (2)

Similarity chains
- (d) was got (= became)
- (e) went walk got = reached)
- (f) lovely dirty
- (g) wash (2) comb brush
- (h) took had (= owned)
- (i) weeks years
- (j) many (2) some
- (k) new different
- (l) speak (2) say
- (m) took-to-bed fell-to-sleep got-up opened-eyes
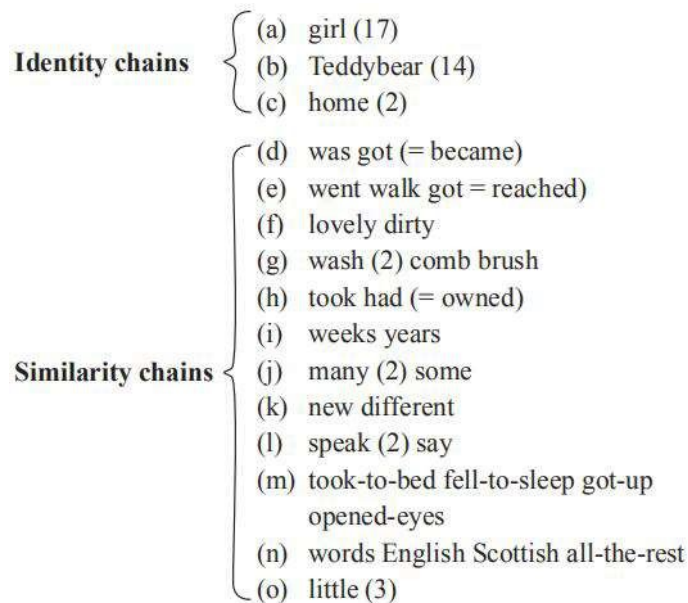- (n) words English Scottish all-the-rest
- (o) little (3)

Figure 3 Cohesive chains in Text 1

We can say, then, that in Text 1, chains　(a) and　(e) interact. A minimum requirement for chain interaction is that at least two members of one chain should stand in the same relation to two members of another chain. This requirement is important for two reasons:

For one thing, the relations that lead to chain interaction are the very ones that exist between the constituents of a clause or of a group, for example, doer, doing; sayer, saying; doing, done-to; or quality, qualified, etc. If a single such relation were considered sufficient for chain interaction, then by definition every member of the chains would interact with some member. This would be tantamount to saying that anything that is a clause or a group is, per se, responsible for coherence. Moreover, there would be no need to differentiate between chain formation and chain interaction; since the former by itself would be a measure of chain interaction. But this is surely wrong since a random list of clauses or groups would not necessarily be coherent; nor does chaining entail coherence.

The second reason is deeper still. The recurrence of a relation between two chains is indicative of two vectors of unity. The first vector of unity is indicated by the semantic similarity that permits members to be part of the same chain; the second vector of unity indicates the semantic similarity that unites at least pairs of members from two chains. The rationale for this is simple to find: in a coherent text one says similar kinds of things about similar phenomena. For example, the girl in
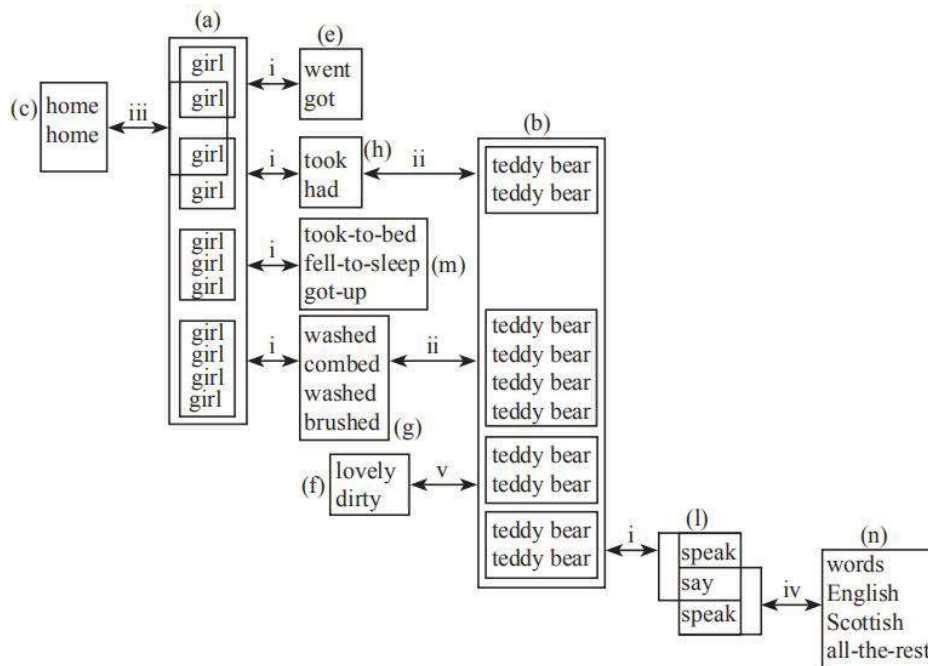

Figure 4 Chain interaction in Text 1

Text 1 does not simply go home, she also gets home; she does not simply fall asleep, she also wakes up, and so on.

When the text is not too long, the chain interaction within it can be visually displayed. This visual display highlights the continuities and the discontinuities in the text. Figure 4 displays the chain interaction in Text 1.

2.5 Cohesive harmony

The total lexical tokens of a text can be classified according to the cohesive chains they enter into:

(1) Relevant tokens: All tokens that enter into identity or similarity chains; these divide into:
      (a) Central tokens: those relevant tokens that interact;
      (b) Non-central tokens: those relevant tokens that do not interact;
(2) Peripheral tokens: All those tokens that do not enter into any kind of chain.

Having established the framework throughout this section, we can now state fairly definitely what the linguistic correlates of variation in coherence will be:

(1)The lower the proportion of the peripheral tokens to the relevant ones, the more coherent the text is likely to be.
(2)The higher the proportion of the central tokens to the non-central ones, the more coherent the text is likely to be.
(3)The fewer the breaks in the picture of interaction, the more coherent the text.

In Figure 4, the entire set of interacting chains is related, with chains (a) and (b) functioning as FOCAL CHAINS, each of which interacts with a large number of other chains.

The three features mentioned above are ordered. The first amounts to saying that the semantic grouping in the text should be such as to establish unequivocally certain definite referential domains. If and when this happens, the majority of the lexical tokens of a text will fall within chains, leaving out but an insignificant few. This is a necessary condition for the second attribute. Texture is thus essential to textual unity, and cohesion is the foundation on which the edifice of coherence is built. Like all foundations, it is necessary but not sufficient by itself.

The second statement amounts to the claim that simply the establishment of the definite referential domains is not enough. Identity and similarity should not be limited to message components alone — such identity and similarity underlie chain formation; the notions of identity and similarity should also be extended to the content of the message as message. In common parlance, when speakers are engaged in the process of creating a coherent text, they stay with the same and similar things long enough to show how similar the states of affairs are in which these same and similar things are implicated.

The third statement claims that the process of creating coherent texts involves an indication of relationships between the things one is ''on about''. The outcome is that a complete break in chain interaction does not take place - transition from one topic to the next is a merging rather than a clear boundary.

The sum of these three phenomena is called COHESIVE HARMONY; and a briefer claim about coherence could be formulated thus: variation in coherence is the function of variation in the cohesive harmony of a text.

It is harmony in more than one respect: it brings together lexical and grammatical cohesive devices, subjecting them to semantic considerations of identity and similarity. This is as it should be; a text, after all, is not a unit of form but of meaning. Secondly, it is harmony because it harmonises the output of two macrofunctions: the textual and the experiential. The output of the textual function are the chains and the interactions; the output of the experiential function at the rank of clause and group is what the interaction is built upon. Thus cohesive harmony is an account of how the two functions find their expression in one significant whole. No doubt, the concept of cohesive harmony can be further refined by bringing in the logical and interpersonal functions into the picture. If this can be done, it will show that even where text is concerned, multifunctionality is a fruitful concept.

3. Analytical procedure

Analytical procedure for cohesive harmony analysis has been implied in the account of the basic concepts. It starts with an analysis of cohesive chains, and then analyzes chain interactions.

Because it brings together the vertical and horizontal dimensions of unity of a text - the kinds of experiential phenomena or semantic domains (realised as lexico-referential chains) present in the text and the grammatical relations into which these enter (represented as interactions between chains) - CHA is able to identify bound-

aries within texts, i.e. the elements of text structure.

The analytical procedure will be illustrated as follows, taken from Cloran, Stu-art- Smith and Young (2007), which gave an example of marking the boundaries of text structure using cohesive harmony analysis.

The first step in cohesive harmony analysis then involves identifying the sense relations between/among the lexically-rendered tokens occurring in different strings or chains of experiential meanings that run through the text. As the concept of chains of experiential meanings implies, the text is considered as it unfolds, i.e. a logogenetic view. Each instance or token of a lexical item in a chain is a component of the message in which it occurs, so chains consist of the same or similar items which are components of messages (clauses). These, then, are the tokens that are relevant in terms of contributing to the cohesion of the text. The text reads as fol-lows:

Table 1 Clause segmentation of Sun Damage text (The text is a modified version of a text from The Australian Women's Weekly, April 1997: 103.)

1	Now that summer is over, you may notice permanent damage to your skin for the first time.
2	Where you have enjoyed many previous summers unscathed,
3	the latest hot season may have been the final straw in damaging the collagen tissue
4	(what I call "the camel's back syndrome").
5	Freckles and blotches are caused by over-stimulation of the pigment cells (melanocytes).
6	These can be treated by bleaching agents used for hair removal or by a ruby laser (at selected dermatologists).
7	The damage to the collagen causes wrinkles
8	and these can be treated at home with glycolic acid and Retin-A
9	(both are available from pharmacies).
10	Chemical peels are also effective.
11	These are done by paramedical technicians in a cosmetic surgeon's or dermatologist's office.
12	More worrying are any sores or blemishes which persist despite three or four weeks of conservative treatment by creams or antibiotics.
13	These need to be tested by a doctor
14	or completely removed
15	and examined under the microscope.

The chains of relevant tokens that run through the Sun Damage text are shown in Figure 5. The lexical sense relations (repetitions, synonyms, antonyms, hyponyms or meronyms) between the tokens are indicated at the bottom of each chain. Chains are labelled for ease of reference. The chains displayed in Figure 5 represent the threads of experiential meanings that run through the text, the (sets of) semantic groupings within the text and thus capture the vertical dimension of the unity of a text.

However, Hasan (1985/89) points out that for the purpose of investigating text connectivity, chain formation alone is not sufficient since a mere listing of items may represent such connectivity, for example, a shopping list or list of ingredients. In analysing the connectedness of a text it is necessary to consider how a component of any individual message relates to the other components of the same message — how chains of related tokens interact with one another. These relations between/among the components of an individual message are the experiential grammatical relations and represent the horizontal dimension of unity of a text. This constitutes the second step in cohesive harmony analysis.

Figure 6 displays the interactions among the chains running through the Sun Damage text.

Now that the vertical and horizontal dimensions of unity of a text has been brought together, we can identify boundaries within texts, i.e. the elements of text structure. This is the third step. The coming and going of chains — their disappearance or appearance as the text unfolds — as represented in the chain interaction diagram (see Figure 6) reveals clearly the three elements in the structure of the Sun Damage text (this structure is not so clearly revealed in the chain formation diagram).

I.Statement of the problem: The focal chain (f) arises here as does the chain that identifies the temporal frame (a). This latter chain disappears when the second stage is introduced.

II.Explanation and recommended action: The chains arising in this stage construct this aspect of the text's field of discourse — chain j (cause), chain k (treatment) and chains l, o and p (chemical preparations to be used in the treatment). Note that the tokens of chains o, l and p interact very closely as either Classifier or Thing. These chains then disappear.

III.Statement of (a further) problem and recommended action: Only the focal chain is continued into this stage but two new chains (r and s) also arise — that related to the investigation of the particular problem.

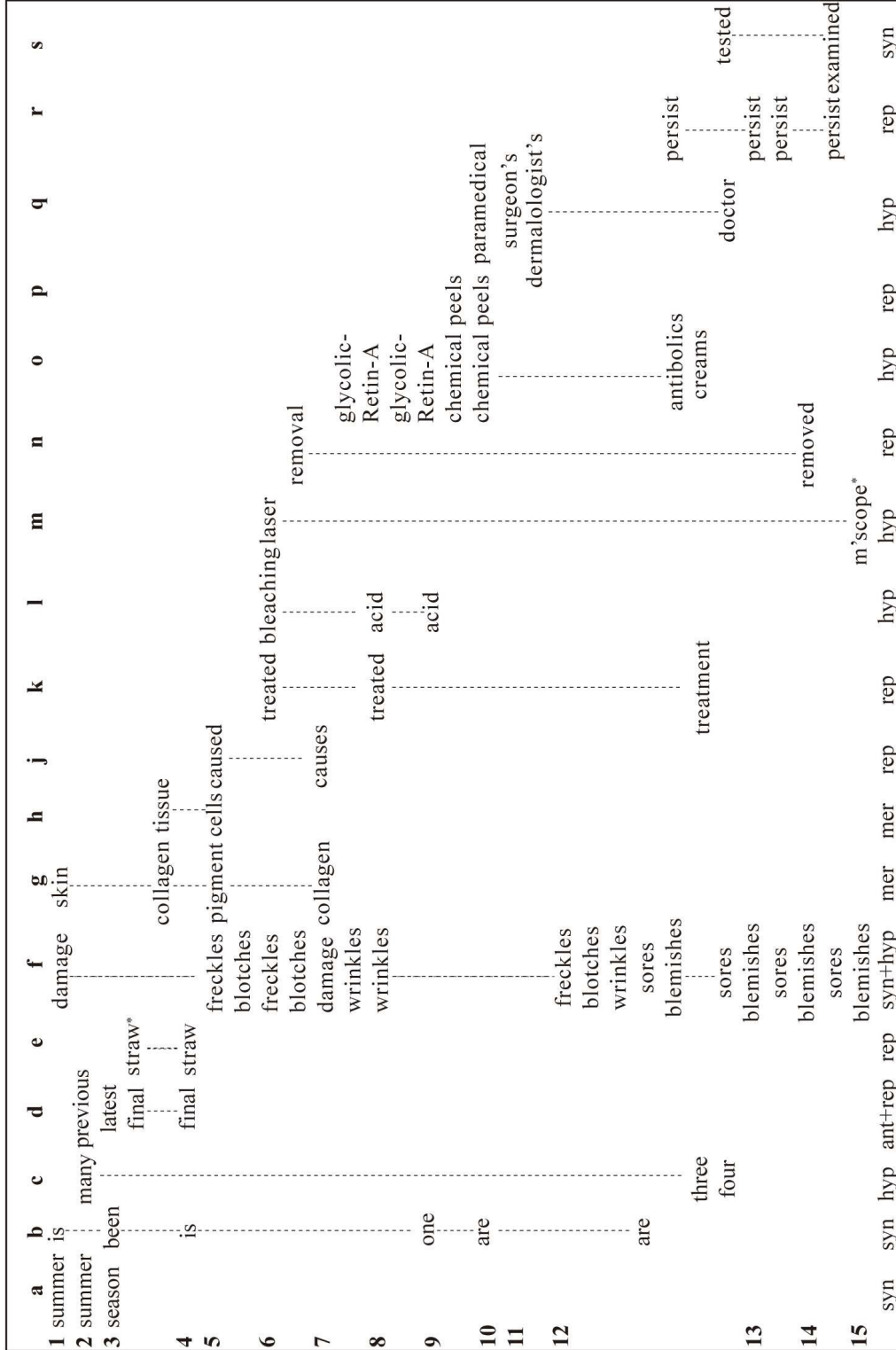The cohesive harmony analysis of the Sun Damage text thus confirms the text

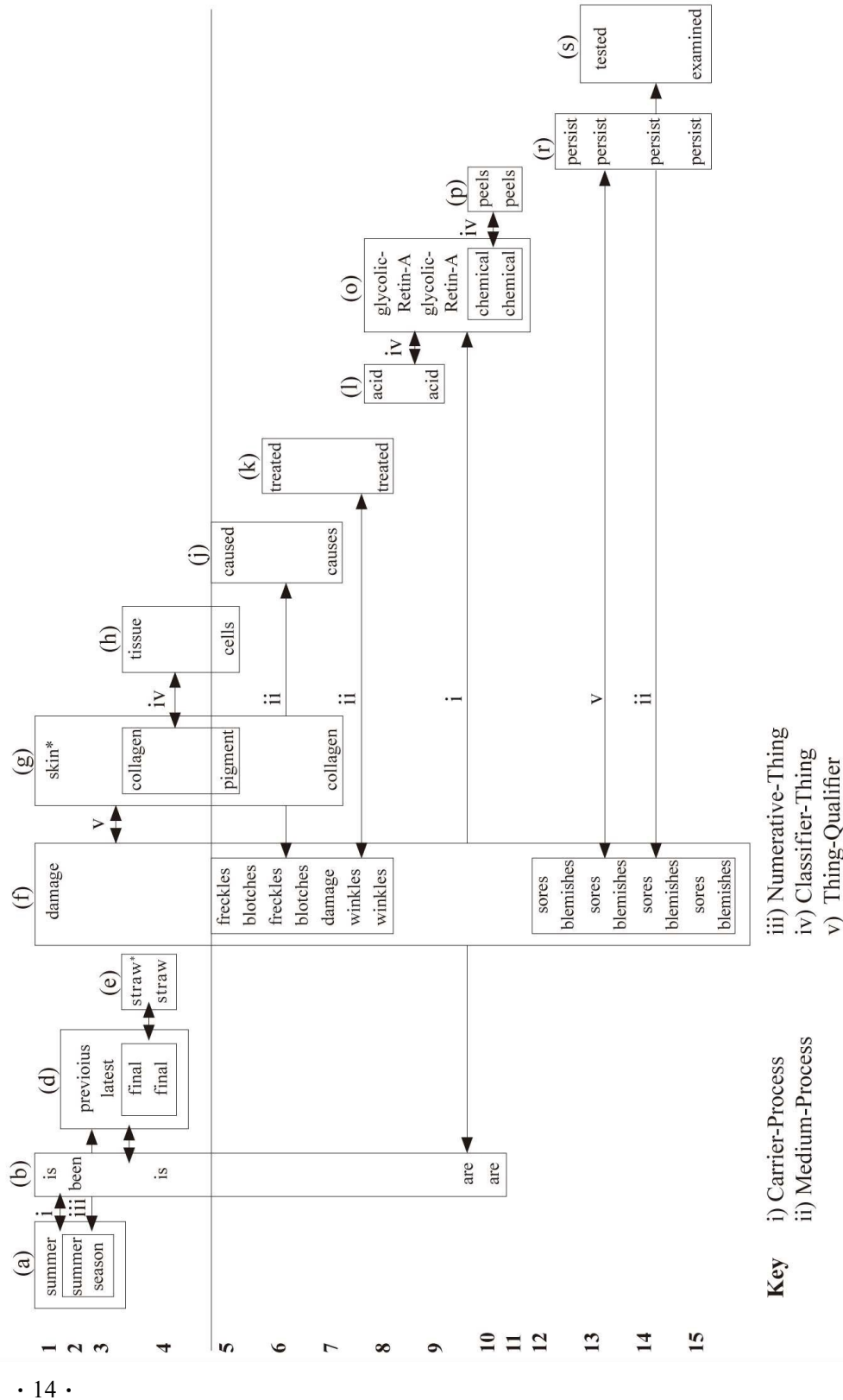Figure 5 lexico-referential chains in the Sun Damage text (relevant tokens)

Figure 5 lexico-referential chains in the Sun Damage text (relevant tokens)

Key   i) Carrier-Process
      ii) Medium-Process
      iii) Numerative-Thing
      iv) Classifier-Thing
      v) Thing-Qualifier

structure previously intuitively identified and provides linguistic evidence support ing the intuition - the (dis)appearance of interacting chains of related lexis.

4. Applications

In addition to identify the boundaries of a text, cohesive harmony analysis can be put to other use. The most frequent areas of application are language teaching and linguistic pathology. CHA has also been used in analysis of identity.

4.1 Language teaching

An example of the application of CHA is Hedberg and Fink (1996). It used CHA to study the use of cohesion in the written stories of normally developing and language learning disabled elementary children. The research indicates that the proportion of words in chains (cohesive density) and chain interaction (cohesive harmony) was consistent across the elementary years. However, children with language-learning disabilities evidenced significantly less cohesive density and cohesive harmony, and they included significantly fewer characters and fewer types of verbal processes in their stories.

4.2 Linguistic pathology

Armstrong (1987, 1991, 1993, 1995, 1997, 2001) conducted a series of analysis of aphasic discourse using CHA and she summarized the research in her 2005 review article.

Lexicogrammatical difficulties also have an effect on overall cohesion or coherence of a text. Cohesion refers to the way in which a text is bound together to form a coherent whole through the use of such grammatical devices as pronouns, demonstra tives, and ellipsis, as well as lexical devices such as repetition, synonymy, antonymy (Halliday & Hasan 1976). Rather than focus on words in isolation, such a system defines the ways in which words within a text must rely on other words for interpretation, and for maintaining topic in a cohesive manner. Hence, rather than looking at word- finding difficulties from a clause level perspective only, the analyst is able to look at the effects of this difficulty on the speaker's ability to mean at the discourse level, i.e. whether s/he is able to construct a coherent text. In both neurogenic and developmental communication disorders, such word-finding diffi culties manifest themselves as pronouns without clear referents, pathological repetition of certain words, and limited variety of lexical choices and

lexical links within texts. The amount to which such phenomena affect listeners was demonstrated by Armstrong (1987) who found a correlation existed between cohesive harmony (a measure of cohesion) and listeners' perceptions of textual coherence. Looking further at the potential contribution of cohesion to coherent discourse, longitudinal data on cohesive links in aphasic discourse has revealed increasing length and variety of cohesive chains as recovery occurs (Armstrong, 1997).

4.3 Identity and identification analysis

Moore (2014) used CHA to probe into the identity issues in an award-winning animal welfare campaign based on mock recipe cards. The study examined how habitual patterns of language make meat-eating and factory farming seem natural, and how certain counter discourses work to expose the seams in such practices. While its genre- bending clearly aims at bypassing reader defenses, the text's real achievement is to combine semantic features whose co-occurrence is normally blocked by the cultural- linguistic system, allowing it to project a sophisticated food identity for readers and construe a social identity for the recipe ''ingredients'' (pigs), realized largely through bizarre cohesive harmony.

5. Conclusion and further development

CHA has proved to be a powerful tool for analyzing text coherence and text structure. Implicit in the cohesive harmony model is a theory of discourse structure as being based in both the syntagmatic and the paradigmatic semantic relations between components of clauses. Coherence arises when the experiential grammatical relations between components of clauses in a text (syntagmatic axis) are mapped onto lexical tokens (paradigmatic axis) that are repeated or semantically related by synonymy (including antonymy, hyponymy and meronymy), these relations being echoed throughout the text (Hasan 1984). Thus CHA captures the fact that in a coherent text 'one says similar kinds of things about similar phenomena' (Hasan 1985/89: 92).

Nevertheless, CHA could be further developed to bring out its potential to the full. For example, it could be considered integrating the cohesive analysis of the interpersonal meaning relations. In addition, and closely related to it, there could be the calibration of cohesive harmony with the analysis of logical relations between clauses.

Works Cited

Armstrong, E.M. Cohesive harmony in aphasic discourse and its significance in listener perception of coherence. In R.H. Brookshire (ed.), Clinical Aphasiology: Conference Proceedings. Minneapolis, MN: BRK Publishers. 1987: 210-15.

Armstrong, E.M. The potential of cohesion analysis in the analysis and treatment of aphasic discourse. Clinical Linguistics and Phonetics 1991, 5: 39-51.

Armstrong, E.M. Aphasia rehabilitation: a sociolinguistic respective. In A. Holland and M. Forbes (eds.), Aphasia Treatment: World Perspectives. San Diego, CA: Singular. 1993:263-90.

Armstrong, E.M. A linguistic approach to the functional skills of aphasic speakers. In C. Code and D. Muller (eds.), Treatment of Aphasia: From Theory to Practice. London: Whurr. 1995: 70-89.

Armstrong, E.M. *A grammatical analysis of aphasic discourse: changes in meaning-making over time*. Unpublished PhD thesis, Macquarie University, Sydney. 1997.

Armstrong, E.M. "Connecting lexical patterns of verb usage with discourse meanings in aphasia," *Aphasiology*. 2001, 15: 1029-46.

Armstrong, E.M. "Language disorder: a functional linguistic perspective," *Clinical Linguistics and Phonetics*, 2005, 19(3): 137-53.

Cloran, C., V. Stuart-Smith & L. Young. Models of discourse. In C.M.I.M. Matthiessen, R. Hasan and J. Webster (eds.), *Continuing Discourse on Language: A Functional Perspective* (Volume 2). London: Equinox. 2007: 647-70.

Halliday, M.A.K. "Categories in the theory of grammar," *Word*. 1961, 17(3): 241-92.

Halliday, M.A.K. & R. Hasan. *Cohesion in English*. London: Longman. 1976.

Hasan, R. Coherence and cohesive harmony. In J. Flood (ed.), Understanding Reading Comprehension. International Reading Association, 1984: 181-219.

Hasan, R. The texture of a text. In M.A.K. Halliday and R. Hasan (eds), *Language, Context, and Text: Aspects of Language in a Social-semiotic Perspective*. Australia: Deakin University Press. 1985: 70-96.

Hedberg, N. & R. Fink. "Cohesive harmony in the written stories of elementary children," *Reading and Writing: An Interdisciplinary Journal*, 1996, 8: 73-86.

Moore, A. "That could be me: identity and identification in discourses about food, meat, and animal welfare," *Linguistics and the Human Sciences*, 2014,9(1): 59-93.