Coherence in Non-technical Summaries of European Central Bank Financial Reports

Aage Hill-Madsen & Helle Dam-Jensen

Abstract Positioned within the research fields of text linguistics and genre studies, this article reports an investigation into how and to what extent coherence is achieved in a specific type of summary writing, viz. non-technical summaries (NTS) of financial/economic reports published by the European Central Bank. The investigation is motivated by the observation that the very nature of summary writing may militate against a sufficient degree of coherence being achieved in such texts. While it is incumbent on the writer of any type of document to ensure coherence in his/her product, this endeavour may be hampered by the fact that summaries are typically required to condense large amounts of information into very limited space. Relying on the Cohesive Harmony Analysis framework provided by Systemic-Functional Linguistics, the investigation found that the texts do succeed in employing cohesive devices in such ways and to such an extent that a clear potential for coherence is present. Nevertheless, the texts fail to exploit the potential in a way that makes them properly coherent.

Keywords coherence, cohesion, expert-lay communication, financial reports, summary writing

1 Introduction

Across EU institutions, it is apparent that great efforts are being made to communicate expert knowledge and information originating in these institutions to the general public. The European Parliament and Council, for example, publish lay summaries of much EU legislation, and in the specific field of health care the EU's medicines regulator, the EMA, publishes summaries for non-expert readers on the outcome of clinical trials preceding authorization of medicinal products. Similarly, the European Central Bank (ECB) publishes numerous leaflets expounding central elements of the Bank's activities and the pillars of its monetary policy.[1] In a number of cases, such knowledge communication consists in the *rewriting* of a source text into a different text type, thus involving a number of changes and adaptations e. g. in register, composition and length. Given the profound changes typically involved, and given the challenges that, accordingly, such textual derivation must be assumed to entail, our research interest in this article concerns the extent to which the rewriting results in (sufficiently) well-functioning *texts*. Textual quality is multifaceted, but in this article we are concerned with one specific aspect only, namely *coherence*, here understood as semantic continuity and connectedness

[1] For summaries of EU legislation, see https://eur-lex.europa.eu/browse/summaries.html.
For EMA summaries of medicinal authorization procedures, see https://www.ema.europa.eu/en/medicines.
For educational ECB publications, see https://www.ecb.europa.eu/explainers/html/index.en.html.

Zitiervorschlag / Citation:
Hill-Madsen, Aage / Dam-Jensen, Helle (2022): „Coherence in Non-technical Summaries of European Central Bank Financial Reports." *Fachsprache. Journal of Professional and Scientific Communication* 44.1–2: 80–101.

(see Section 3), which is an underresearched aspect of knowledge communication texts.[2] It is, nevertheless, one that must be deemed crucial to textual quality in genres aimed at non-expert readerships (or any type of readership, for that matter), considering that lack of coherence is a clear obstacle to successful communication (Hill-Madsen 2019; see Section 2 below).

A particular subset of rewritten knowledge-communication texts that must be deemed especially prone to problems of insufficient coherence are *summaries* of much longer texts, given that summaries are products of both condensation and selection of information (Hidi/Anderson 1986: 473). One example of a very condensed type of summary is featured in one of the opening sections of economic and financial reports published by the European Central Bank, viz. the non-technical summary (NTS), aimed at a non-expert readership (see Section 4). The source texts behind the NTSs are around 45 standard pages long on average and their summaries around two and a half standard pages, making the latter very compact texts that include only the most salient points.[3] Since the high degree of condensation involved in the summarizing may be assumed to be at least *potentially* detrimental to the creation of target texts with a sufficient level of semantic continuity and connectedness, our research aim is to explore whether such connectedness can be adequately attained in these texts, and our research question is thus: What is the nature and degree of coherence in NTSs of ECB economic/financial reports?

For analytical methodology, the investigation relies on the framework proposed by Halliday/Hasan (1976) and Hasan (1984, 1989), which provides a rich theoretical foundation for analyzing the use of various cohesive devices to achieve coherence throughout a text (Hasan 1984: 187 f.).

2 Summaries

Summary writing is an "intermediate" text-production type, in the sense that it involves a shift in text type from the source text to the rewritten product (Jakobsen 2005: 176, Dam-Jensen/Heine 2013: 92). As noted in the Introduction, the shift is centrally a matter of condensation, the purpose being that of representing the core information of a source text in a shorter form.

Producing a condensed version of a source text involves a number of rewriting processes, such as deletion of inessential information, generalization of detailed information, and integration of details into topic sentences (Hidi/Anderson 1986: 480, Mani 2006: 277), corresponding, to some extent, to van Dijk/Kintsch's (1978) macro-rules of construction, deletion, and generalization. These processes are used to varying degrees, depending on the difference in length between the source text and the summary and on the type of summary to be produced. As sub-types of summaries, Mani (2006: 274) mentions extracts and abstracts, each covering further subtypes, depending on purpose. As an example of rewriting operations in a

[2] A parallel aspect of textual quality in knowledge-communication texts for the general public is the question of 'lay-friendliness', i. e. whether a satisfactory degree of popularization is consistently attained in the texts. However, while still an important issue, this is an aspect that has already received considerable attention in LSP research, especially within lay-oriented health communication (cf. e. g. Hill-Madsen/Pilegaard 2019, Askehave/Zethsen 2000).

[3] As products of rewriting, a genre like the NTS may lend itself to a range of inquiries regarding the relation between source and target texts, e. g. how and to what extent the different sections of the source text are represented in the target text. In this article, however, the focus is on the target texts in isolation.

specific summary type, some measure of copy-pasting is typically required in extracts (Mani 2006: 274).[4]

Whatever the specific type, summaries are subject to two basic requirements: coherence and informativeness (Mani 2006: 275). A sufficient degree of informativeness is attained by ensuring that the summary is faithful to the source text and that it is relevant to its purpose and users (Mani 2006: 275). Further, we hold that informativeness is contingent on coherence, given that an incoherent text is bound to jeopardize the meaningful relay of information. The very nature of coherence as a textual phenomenon, and how it is achieved, will be theorized in Section 3 below.

3 Coherence and cohesive devices

In order for a text to be meaningful, its parts must be connected structurally as well as semantically. Textual connectedness is often accounted for in terms of cohesion and coherence, which are well-known concepts within text linguistics, but far from unambiguous ones. With the purpose of contextualizing the two phenomena, we will start by presenting a short historical review of various interpretations.

3.1 Theoretical approaches to cohesion and coherence

In the realm of structural linguistics, the approach to text structure was descriptive and text was defined as a unit larger than the sentence. Studies of textual connectivity amounted to classifying text structure (de Beaugrande/Dressler 1981/2016: 23). Harweg (1968) was one of the first contributions to a "larger-scale study of text organisations" (de Beaugrande/Dressler 1981/2016: 22) and, in his framework, a text is created by way of substitution, in the sense that sentences are connected to form a text "by an uninterrupted chain of syntagmatic substitutions" (Harweg 1968: 148).

In a semiotic perspective, Greimas (1966) opened the scope of coherence to include semantic isotopy, i. e. the semantic relations between lexemes, at any level of the text. According to Greimas/Courtés (1979: 197–199), the semantic coherence of a text is established through the repetition of (lexical) semantic elements (*classèmes*) and the recurrence of semantic categories (*catégories sémiques*). Referring to semantic relations, semantic isotopy thus broadened the concept of coherence beyond a merely structural understanding.

Later developments in linguistics, such as Cognitive Linguistics and Functional Linguistics, similarly moved beyond a merely structural understanding of the mechanisms behind textual coherence. In Cognitive Linguistics, textual relations are established by the reader as a mental representation based on linguistic input. In this sense, a text is coherent as a result of the representations that interlocutors establish of the segments (Sanders/Spooren/Noordman 1993: 114), which leads to the creation of a "mental textual world" (Menzel/Lapshinova-Koltunski/Kunz 2017: 1). The French linguist Michel Charolles thus explains coherence as the result of a dynamic interpretive process whereby the receiver of discourse establishes meaning relations incrementally: "Confronté a une séquence d'énoncés produits à la suite, le destinataire ne peut en effet que chercher à établir des relations entre ces énoncés, vu que,

[4] It should be noted that Mani (2006) is a study of automatic summarizing, but the criteria and characteristics mentioned here are applicable to human summarizing also.

précisément, ils sont énoncés à la suite." [When confronted with a sequence of consecutive utterances, the receiver will try to establish relations between the utterances, precisely because they follow each other. (*Our translation*)] (Charolles 1997: 3) Cohesive markers are seen as the instructional input for this process to run (cf. Givón 1995).

In both Cognitive and Functional Linguistics, textual connectivity is generally explained in terms of a distinction between cohesion and coherence, in that both linguistic paradigms conceive of coherence as a semantic quality signalled by linguistic markers, i. e. cohesive devices. The point of contention between the two paradigms resides in the very understanding of the nature and 'location' of semantics, and hence coherence: whether meaning is a mental phenomenon (the cognitive approach) or a text-immanent quality (the systemic-functional approach; cf. Halliday/Matthiessen 1999: 416). This debate is, however, far beyond the scope of the present article. What matters to our purposes is the point of concurrence between the paradigms: both assume that textual coherence is a matter of degree, depending on a more or less elaborate and successful use of cohesive devices.[5] Accordingly, it makes sense to investigate how and to what extent texts are cohesive, which is the particular concern of the present investigation. For the appraisal of cohesiveness, we find the tools offered by Systemic-Functional Linguistics particularly useful, which is why certain aspects of this framework will be briefly introduced below.

3.2 Cohesive devices: types and functions

In the text-linguistic literature (cognitive as well as functionalist) (cf. e. g. de Beaugrande/Dressler 1981/2016, Hasan 1989, Martin 1992, Taboada 2004, Sanders/Sanders 2006, Helder 2011/2015, Halliday/Matthiessen 2014), the following cohesive devices are generally recognized: (1) conjunction, i. e. the use of connectors, (2) reference, (3) substitution and ellipsis, and (4) semantic isotopy (or lexical cohesion).[6] Apart from these devices, a number of studies (e. g. Halliday 1967, Firbas 1974, Chafe 1976, Fries 1983, Lambrecht 1994, Gundel/Fretheim 2004, Büring 2007, Krifka 2008, Halliday/Matthiessen 2014) stress (5) information structure and thematic structure as important elements in the progression of text from sentence to sentence.

Below, those of the concepts that are significant to our purposes will be further commented on. Though some or all of them are likely to be well-known text-linguistic concepts, we nevertheless deem it necessary to provide some degree of definition (from the Systemic-Functional framework) as well as a further comment for the purpose of operationalizing the concepts for our specific analytical aims (see Section 4).

(i) Conjunction:
The use of connectors is the linguistic manifestation of the logical relations between propositions, for example relations of addition, causality, and opposition (Halliday/Hasan 1976: 226–273, Martin 1992: 159–270, Halliday/Matthiessen 2014: 611–622). In many cases, though, the logical connection is not made explicit by means of connectors, but should be inferred from

[5] The relation between coherence and cohesion as defined here is our interpretation of the systemic-functional position. It should be noted, however, that Halliday and Hasan do not always make a sufficiently clear distinction between the two concepts.

[6] It is important to note that not all authors mention all the listed phenomena.

the content in combination with extralinguistic context. In the analyses, however, we will follow Halliday/Matthiessen (2014: 622) in only recognizing logical connections that are explicitly coded by means of connectors.

(ii) Reference:
In accordance with Halliday/Hasan (1976: 38) and Halliday/Matthiessen (2014: 605 f.), we understand *reference* as the links established by referring expressions, primarily in the form of demonstrative and personal pronouns. What is significant to our purposes is the fact that by maintaining referents, the use of pronouns creates chains running through a text, e. g. *the world economy … it … it*, or *enterprises … these … they* (for the further significance of cohesive chains, see Sections 3.3 and 3.4).

(iii) Information structure (Given-New) and thematic structure (Theme-Rheme):
Theme-Rheme structure in particular is a well-known concept within text linguistics. Theme, which occurs in clause-initial position, indicates the speaker's 'topic' or point of departure for his/her message, whereas Rheme provides the actual content of the message (Halliday/Matthiessen 2014: 89). Thematic structure is speaker-oriented, whereas information structure is listener-oriented, accounting for the distribution between 'given' (i. e. already known) information and 'new' information. As elements in a functional structure, the two constituents (Given and New) are only separated from each other phonologically, viz. by means of intonation (Halliday/Matthiessen 2014: 114–119), and are thus not explicitly cued in writing. Yet, Martin (1992) points out that in English, NPs are in fact grammatically coded as either 'given' or 'new'. In Martin's words, "[b]asically, indefinite nominal groups [*i. e. NPs*] code the identity of the participant being realised as not recoverable [*i. e. as new information*], whereas pronouns, demonstratives, the definite article and proper names signal that the participant's identity is in some sense known" (Martin 1992: 92). Our analysis of information structure in Section 5.1 will be informed by Martin's distinctions.

What is significant to our purposes is how the two different structures interact to create textual connectivity. As already emphasized, Given-New and Theme-Rheme are independent of each other (cf. Fries 1983), yet the default interaction between them consists in the coincidence of given information with Theme (Halliday/Matthiessen 2014: 120). In other words, a clause or a sentence can usually be expected to take its point of departure (Theme) in information that is already known to the recipient (Given). Such overlap between Theme and Given is an effective way of linking the content of one sentence with preceding context (or text-external context, in the case of thematic NPs with exophoric reference) (Hill-Madsen 2019: 164). In the analyses in Section 5.1, therefore, a particular part of the investigation will be concerned with exploring the extent to which the sampled texts observe this principle of mapping Given information onto the Theme of each sentence.

(iv) Semantic isotopy (lexical cohesion):
As mentioned above, the concept of semantic isotopy was introduced into text analysis by semiotics to account for the semantic unity between lexemes of a text. Semantic isotopy is here understood as corresponding, at a general level, to the phenomenon of lexical cohesion first introduced by Halliday/Hasan (1976). In their conception, lexical cohesion manifests itself in semantic ties between lexical items occurring in different sentences (but sometimes also within the same sentence). Ties between lexemes rely on the well-known types of sense relations, viz. synonymy, antonymy/opposites, hyponymy/hyperonymy, co-hyponymy, meronymy/hol-

onymy, and co-meronymy (Hasan 1989: 80 f.). Apart from ties based on sense relations, one further, and frequent, type is sheer repetition of the same lexeme (Hasan 1989: 81). Examples of lexical ties are provided in the text below (an excerpt from a style guide published by the Plain Language Movement in the US (Plainlanguage.gov 2011):

Example 1:

> [1] *Short sections break up material so it appears easier to comprehend.* [2] *Long, dense sections with no white space are visually unappealing, and give the impression your document is difficult to understand.* [3] *Short sections appear easier to comprehend, and help you organize your document more effectively.* [4] *Short sections also give you more opportunity to insert informative headings in your material.* [5] *Remember that boldface section headings give your reader the best roadmap to your document.* [6] *Long sections are impossible to summarize meaningfully in a heading.* [7] *When you write short sections, each heading can give the reader information about the entire contents of the section.* (p. 15) [sentence numbering inserted by us, AH-M & HD-J]

The following lexical ties are instantiated in the excerpt:

- Synonymy: *comprehend – understand*, *appear – give the impression*
- Antonymy/opposites: *easier – difficult*, *short – long*, *read*(*er*) *– write*
- Meronymy/holonymy: *document – sections – headings*
- Repetitions: *short – short*, *material – material*, *sections – sections*, *headings – heading*, *you – your*, *informative – information*, *is – are*, *give – give – give*, *reader – reader*

Lexical ties give rise to multi-token chains running 'vertically' through the text (Hill-Madsen 2019: 161). Thus, in Example 1, the altogether five items *short* [sentence 1] – *long* [2] – *short* [3] – *short* [4] – *long* [6] form a chain (or 'semantic isotopy') whose links consist in antonymy between [1] and [2], [2] and [3], and [4] and [6], and in repetition between [3] and [4]. The example illustrates that there may well be variation in the type of ties that make up a chain. Semantically, a chain is the lexical manifestation of a subtheme in the text, such as 'length'[7] in the case of the five-item *short – long* isotopy. In the same way, the meronymic relations between *document*, *sections* and *headings* create a 'document' isotopy. Other isotopies in Example 1 are 'material', 'comprehension' (*comprehend – understand*), 'you', 'difficulty' (*easier – easier – difficult*), 'information', 'giving', 'being' (*is – are*), 'appearing' (*appear – give the impression*), and 'reading/writing' (*reader – reader – write*).

3.3 The principles of Cohesive Harmony

According to Hasan (1989: 89–91), the importance of chains (both lexical and referential ones) resides in their contribution to connectivity: The higher the number of message components[8] entering into chains, the stronger the connectivity of the text as a whole. However, whereas the organization of message components into chains is a necessary condition for textual connectivity, it is not a sufficient one. While chains help create thematic (i. e. 'topic-related') connec-

[7] Henceforth, one-word (or, in very few cases, multi-word) names in single inverted commas will be used to refer to individual chains. In each case, the name chosen is one that best captures the semantic content or theme of the chain. Lexical items representing the chain are italicized.

[8] The term *message components* will be used henceforth as an umbrella term for lexical and referential items.

tions between different *sentences*, the semantic *themes* (manifested in chains) also need to be interlinked. Hasan (1989: 91) refers to such linkage as *chain interaction*. In sum, connectivity in a text at large is achieved if:

- The text contains chains
- The chains are interlinked

The combination of these two principles is what Hasan (1989) has dubbed Cohesive Harmony. Linkage between chains is achieved 'horizontally' via syntagmatic relations between members of different chains, which is also manifested in Example 1 above. Chain interaction occurs, e. g., between the two isotopies 'length' and 'documents' in sentence [1] via the interrelation of the two lexical items *short* and *section* as Premodifier and Head (*short sections*). In the same sentence, a member of the 'appearing' isotopy (*appears*) and a token from the 'difficulty' chain (*easier*) interact as Verb and Subject Complement (*appears easier to comprehend*). The 'difficulty' chain (*easier*) and the 'comprehension' chain (*comprehend*) interact as Head + Postmodifier (*easier to comprehend*).

However, a single occurrence of interrelation between two chains does not suffice to establish proper linkage between them. The reason is the obvious one that such interrelation is present in most cases anyway: Since all message components are syntagmatically related to other components, this means that all chains are likely to interact with other chains (Hasan 1989: 91), at least if a sufficient number of the message components of a text enter into chains. Instead, in order for chains to be properly linked, at least two points of similar grammatical interaction between two chains is necessary. Such recurrence of the same syntagmatic relation between two items from different chains is, in the words of Hasan (1989: 94), the grammatical manifestation of "say[ing] similar kinds of things about similar phenomena", which is what one does in a coherent text (Hasan 1989: 94). In less popular terms, coherence centrally consists in the syntagmatically created integration of semantic subthemes. As proposed by Hill-Madsen (2019), we will call such recurrent chain interactions *chain interaction series* (CIS).

In the short text in Example 1 above, all chains are in fact linked with other chains in CISs via the recurrence of similar syntagmatic configurations between members of different chains. Altogether 7 different CISs can be identified:

- The 'length' + 'documents' isotopies are connected as Premodifier and Head in sentences [2] (*long sections*), [3] (*short sections*), [4] (*short sections*) and [6] (*long sections*).
- The 'appearing' + 'difficulty' isotopies are connected as Verb and Subject Complement in sentences [1] (*appears easier*) and [3] (*appears easier*).
- 'difficulty' + 'comprehension' isotopies are connected as Head and Postmodifier in sentences [1] (*easier to comprehend*) and [3] (*easier to comprehend*).
- 'documents' + 'being' are connected as Subject and Verb in sentences [2] and [6].
- 'documents' + 'giving' are connected as Subject and Verb in sentences [4] and [7].
- 'you' + 'documents' are connected as Determiner and Head in sentences [2] and [3].
- 'giving' + 'reading/writing' are connected as Verb and Indirect Object in sentences [5] (*give your reader*) and [7] (*give the reader*).

Thus, all chains except the 'materials' isotopy participate in CISs in Example 1. Furthermore, most, but not all, CISs are connected with other CISs, as is apparent from the graphic illustration in Figure 1 below:

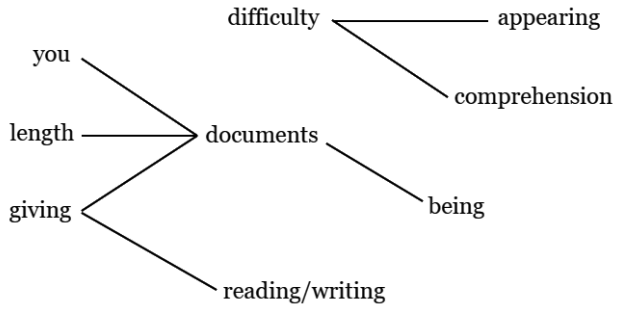

Figure 1: The network of CISs in Example 1

Figure 1 graphically represents the network of interacting chains in Example 1. The graph illustrates that there is an interactive gap (Hasan 1984: 201) in the 'mesh', disconnecting three of the chains (the 'difficulty', 'appearing' and 'comprehension' chains) from the rest of the network. Hasan's point (1989: 93) is that in a coherent text, such interactive gaps are few or even non-existent (see further below).

3.4 Operationalization of the principles of Cohesive Harmony

The principles of Cohesive Harmony set out in the previous section may be summed up in four terms that account for the degree of textual connectedness: *relevant tokens*, *peripheral tokens*, *central tokens* and *non-central tokens*.[9]

Relevant tokens are those that participate in chains, whereas *peripheral* ones are those that do not. In Example 1 (in Section 3.2 above), items such as *visually*, *unappealing* and *organize* were instances of peripheral tokens, falling outside isotopies.

Central tokens, in turn, cover the subset of *relevant tokens* that participate not only in chains, but also in CISs, whereas *non-central ones* are those *relevant tokens* that do not enter into CISs (Hasan 1989: 93).

The four concepts in fact provide the basis for quantification since, according to Hasan (1989: 93), the semantic connectivity of a text (at least those aspects to do with referential and lexical chains) is quantifiable as:[10]

1. The ratio of relevant to peripheral tokens. The higher this ratio, the stronger the textual connectivity (is likely to be).
2. The ratio of central to non-central tokens. The higher this ratio, the stronger the connectivity (is likely to be).

[9] *Tokens*, which is Hasan's (1989: 93 *et passim*) term, should here be understood as synonymous with *message components*, encompassing lexical, referential, substitutional and elliptical items. In connection with the calculation of central, non-central and peripheral tokens, Hasan recommends a procedure termed *lexical rendering* (Hasan 1989: 87), whereby referential, substitutional and elliptical items are in the individual case interpreted and registered as the corresponding lexical item they refer back to or replace. In connection with an anaphoric pronoun such as *they* referring back to, e. g., *enterprises*, both would be registered as tokens of the lexical item *enterprise*. This procedure of *lexical rendering* is followed in the present investigation.

[10] This outline is based on the account in Hill-Madsen (2019: 163).

3. The number of interactive gaps. The lower the number of interactive gaps in the network of interlinked CISs, the greater the connectivity of the text.

The relevant-peripheral ratio (1) reflects a text's degree of integration with regard to subject matter: If a text features a low percentage of peripheral tokens, it thereby has few items on the periphery of the text's thematic focus. However, although a high relevant-peripheral ratio is a *sine qua non* for a proper degree of textual connectivity, it is not sufficient. A high central-non-central ratio (2) is also necessary, meaning that the sub-themes of a text (manifested in chains) need to be extensively linked with each other. A final requirement is that the development of the subject matter from one sub-theme to another must be a connected, uninterrupted process (Hasan 1989: 94). Interactive gaps (3) represent such interruptions, which is why in a sufficiently coherent text the number of such gaps is low.

In sum, given the framework's mixture of qualitative and quantitative elements, we find the principles of Cohesive Harmony analysis particularly suitable for the investigation of thematic (i. e. 'topic-related') connectivity in a text. Qualitatively, the framework provides tools for uncovering the specific linguistic resources employed to establish connectivity in a given text, while at the same time it enables the connectivity to be quantified.

4 Corpus

The corpus on which the investigation is based consists of six non-technical summaries (NTSs) sampled from the publications section of the European Central Bank's website.[11] Given that the summaries are not only shortened, but also 'non-technical' versions of their source texts, the target readership must be assumed to be the general public.[12] Thematically, the ECB's publications are concerned with economic, financial and statistical subject matters related to the Bank's central task of formulating and implementing monetary policy for the Eurozone. In the ECB publications, the NTS occurs as the second section of the text as a whole, following immediately after a very short abstract and preceding the report proper. In terms of condensation, the NTSs are 6 % the length of their source texts on average.

The six texts in the corpus[13] were selected through *purposeful random sampling* (Patton 2002: 240 f.). Random sampling from the whole population of ECB publications was opted for since the aim was to obtain a broad sample of NTSs within the field of finance and economics, enabling conclusions about the nature and possibility of coherence in economic/financial NTSs at large, rather than within a narrower population defined e. g. by a more specific sub-field. The random sampling has thus resulted in a broad range of subject matters being represented, covering fiscal policies (Text C), developments in global trade (Text E), specific types of monetary-policy instruments and money-market operations (Texts D and F), and certain statistically oriented subjects relating to economic forecasting and government debt and budget deficits (Texts A and B). The other aspect of the sampling method (the

11 See https://www.ecb.europa.eu/pub/html/index.en.html.

12 Unfortunately, no guidelines specific to the ECB publication NTSs appear to exist, but their intended target readership may be induced by comparison with a parallel genre, viz. that of Environmental Impact Assessments NTSs. This latter genre is required to "make the key issues and findings of the environmental statement accessible and easily understood by the general public." (Institute of Environmental Management & Assessment 2012).

13 A list with the full bibliographical details of all six sampled texts is provided in Appendix 1.

purposeful element) consisted in the inclusion of a particular selection criterion, which was length. Since the ECB NTSs vary in length (between approximately 1.5 and 4 standard pages), the six texts in the corpus were sampled so as to reflect this variation, with the shortest text at 1.4 pages and the longest at 3.8. The purpose of this variation was to investigate whether there are reasons to hypothesize that the degree of connectivity correlates with text length (see Section 5.6).

5 Qualitative and quantitative analyses

In accordance with the theoretical account of cohesive devices and textual connectivity (Section 3), the analyses proceeded along six distinct steps, of which the first five are qualitative and the sixth is quantitative. Steps 1–2 are concerned with the continuity between adjacent sentences, and steps 3–6 with chain-related connectivity:

- Steps 1 and 2: All sentences in each text were analyzed for the presence of conjunctive items and for the interaction of Theme-Rheme and Given-New patterning. The two steps will be documented in Section 5.1.
- Step 3: All cohesive chains in each text were identified through analysis of cohesive relations (referential and lexical). For documentation of this part, see Section 5.2, where the analysis of a single sample text will be presented.
- Step 4: Those clauses and phrases featuring neighboring members of chains were syntactically analyzed with a view to identifying CISs. This part of the analysis is exemplified in Section 5.3 by means of the same sample text.
- Step 5: For each text, a map was created showing the interaction as well as the interactive 'gaps' between chains. A representative example of one of these maps will be shown in Section 5.4.
- Step 6: The results from the chain and CIS analysis were quantified, and the ratios introduced in Section 3.4 (relevant-peripheral and central-non-central ratio) were calculated. The quantitative results are presented in Section 5.5.

5.1 Analytical steps 1 and 2: Sentence-to-sentence continuity

For all of the six sampled texts, the analysis reveals that the sentence-to-sentence continuity must be considered satisfactory. This is because in the vast majority of cases, the use of connectors as well as the interaction between Theme-Rheme and Given-New is such that adjacent sentences are properly connected by either or both of these two cohesive means. In the following, a single representative and illustrative text (listed as Text E in Appendix 1 and reproduced in full in Appendix 2)[14] will be selected for exemplification of the qualitative results of the analyses.

Use of connectors: Conjunction plays only a minor part in establishing sentence-to-sentence continuity in the corpus at large. In Text E, only five instances of connectors occur, viz. *as a result* [sentence 4], *moreover* [sentence 26], *for instance* [sentence 31], *yet* [sentence 33],

[14] It should be noted that one respect in which Text E is atypical is the fact that it includes a table. Since the table is an instance of multimodal semiosis, and one, moreover, that is not integrated with the summary text, it has been ignored in the analysis and is not included in Appendix 1. In the present investigation, coherence is only investigated as a feature of verbal-language *text*.

a*s such* [sentence 34]. In all five cases, the connector occurs sentence-initially, as is customary for these items (cf. Halliday/Matthiessen 2014: 109).

Interaction of Theme-Rheme and Given-New: Given the limited use of connectors, sentence-to-sentence continuity is mainly established through the choice of thematic material. In Text E, Theme generally conflates with 'given' information, typically realized by a definite NP (cf. Section 3.2) referring back to an item in the immediately preceding sentence (in a few cases the penultimate sentence) or by an NP with exophoric reference. An example of the former, with the Theme in the latter of two adjacent sentences underlined, is:

Example 2:

> [9] *One source of change arises from compositional effects, such as the shift of growth in trade and economic activity towards economies with lower trade intensity, and changes in the composition of aggregate demand factors towards less trade-intensive components.* [10] *<u>These shifts</u> are not necessarily structural and could reverse in part over the medium term.*

In sentence [10], the thematic NP *These shifts* refers back to the economic and trade-related shifts and changes mentioned in sentence [9].

An example of an exophoric Theme is:

Example 3:

> [27] *<u>This paper</u> concludes that the recent weakness in the global trade-income relationship constitutes a "new normal" …*

This paper is an exophoric NP, in so far as reference is being made to the constitutive element of the communicative event itself. Exophoric Themes (of which two-three instances are present in all of the sampled texts, either as *this paper* or *this report*) may be considered a borderline case of connectivity. This is because they do not link the sentence where they occur with a preceding one, and are thus, strictly speaking, not a cohesive device (cf. Halliday/Hasan 1976: 13). Nevertheless, references to the paper in itself do serve to anchor the sentence in relevant context, albeit this context is the communicative situation at large and not preceding *text*.

5.2 Analytical step 3: Identification of cohesive chains

Although the particular nature of the cohesive links in the corpus is not *per se* of interest to the main purposes of this paper, a couple of general characteristics appearing from the purely cohesive analysis (i. e. the identification of cohesive chains) may be noted. Firstly, it turns out that the chains identified in the corpus at large are mainly lexical ones and only in a small minority of cases referential ones. Secondly, the analysis revealed that the links between the lexical parts of the chains mostly consist in mere repetition of the same lexical item (though typically in different derivational forms). The second-most (but much less) frequent type of link is synonymy, with hyponymy, co-hyponymy and antonymy occurring only in few cases. In Text E an example of a chain exhibiting synonymy and antonymy apart from repetition is one concerned with the subtheme of 'globality'. This chain consists in 18 instances of the adjectival form of *globe*, 3 instances of its near-synonym *world*, one instance of the adjective *overall* (which in the particular case must be interpreted as being synonymous with *global*), and one instance of the antonymic *local*. Similarly, a chain having to do with 'increase/decrease' consists in a number of synonyms related to the sense 'decrease', e. g. *decelerated, declined* and *diminishing*,

and two synonymous items (*increasing* and *rise*) realizing the opposite sense 'increase'. (For an overview of all chains identified in Text E, see Appendix 3.)

5.3 Analytical step 4: Identification of Chain Interaction Series

From Text E, two examples have been selected for illustration of chain interaction:

Example 4:

> [9] *One **source** of **change** arises from **compositional effects**, such as the shift of **growth** in **trade** and economic activity towards economies with **lower trade intensity**, and changes in the composition of aggregate demand factors towards less trade-intensive components.*

In Example 4, all the underlined items are members of chains, and the bolded ones are those that enter into Chain Interaction Series. Thus, 'source' and 'change' are part of a CIS in which the two chains interact via a Head + Postmodifier relation; 'components' (represented by the item *compositional*) and 'effects' are part of a CIS consisting in a Premodifier + Head relation; 'growth' and 'trade' interact via a Head + Postmodifier relation; and 'lowness' (instantiated in *lower*) and 'trade intensity' interact in a Premodifier + Head relation also.

Example 5:

> [24] *The third **channel reflects diminishing marginal support** from **financial deepening** to facilitate export capacity.*

In Example 5, four different CISs are represented, viz. 'channel' + 'reflecting', 'reflecting' + 'support', 'increase/decrease' + 'support' and 'support' + 'financial deepening'. Of these four, the connection between the two chains 'channel' and 'reflecting', on the one hand, and between 'reflecting' and 'support', on the other, are some of the relatively few instances of CISs based on clause-level associations. In the case of the 'channel' + 'reflecting' CIS, the interaction consists in a configuration between Subject (*channel*) and Verb (*reflects*), and, in the case of the 'reflecting' + 'support' CIS, between Verb and Subject Complement (*support*). The 'increase/decrease' + 'support' CIS (realized in the items *diminishing* and *support*), on the other hand, represents a phrase-level connection, viz. Premodifier + Head. The 'support' + 'financial deepening' CIS, too, is a phrase-level configuration, consisting in Head + Postmodifier.

In the corpus at large,[15] the vast majority of grammatical connections tying two chains together are in fact phrase-level ones, especially the type consisting in Premodifier + Head, whereas clause-level configurations are relatively rare. Given the particular syntactic characteristics of the texts, this is hardly surprising: Without exception, all the texts are syntactically characteristic of written-style language, exhibiting a high complexity at phrase level, but low complexity at sentence level (cf. Halliday 1989, 1987/2002). Example 6 illustrates the phenomenon:

Example 6:

> [20] *The first reflects waning support from factors that had previously contributed to global trade outpacing global output growth, including lower transportation costs, the removal of trade barriers through lower tariffs and the growing adherence to global trade agreements through the increase in WTO membership.*

[15] For a complete list of CISs identified in Text E, including information about the type of grammatical relation tying the chains together in each case, see Appendix 4.

The underlined strings in Example 6 are good examples of a number of long and heavily 'stacked' noun phrases with heavy (post-)modification featuring phrases within larger phrases. In the sentence as a whole, on the other hand, there are only two clauses, with *the first reflects* constituting the Subject and Verb of the main clause, and *that had previously contributed to …* serving as a modifying clause within the noun phrase whose Head is *factors*. This means that the potential for phrase-level connections between two chains far outweighs the potential for clause-level ones. It may be noted that this imbalance between the number of clause-level and phrase-level constructions is centrally to be explained by the phenomenon of nominalization. An important consequence of nominalization is that configurations of semantic units that are usually expressed at clause level (Halliday/Matthiessen 1999: 227) are 'downranked' to phrase level, in Halliday/Matthiessen's (2014: 614 *et passim*) parlance. Thus, the noun phrases *waning support from …*, *global output growth*, *lower transportation costs*, *the removal of trade barriers*, *the growing adherence to …* and *the increase in WTO membership* may all be 'unpacked' (Halliday/Matthiessen 1999: 255) at clause level to reveal configurations of 'agents' and 'processes/actions'. Clausal versions of the above NPs would thus be *support from … is waning, global output is growing, transportation costs less,* [*governments*] *have removed trade barriers*, [*governments*] *increasingly adhere to …* and *WTO membership is increasing*. Nominalizations such as *support*, *output* and *membership* might of course be even further 'unpacked', which, however, will not be done here.

5.4 Graphical illustration of chain interaction

Figure 2 below is a diagrammatic illustration of all interactions (and lack of interactions) between cohesive chains in Text E, which has once again been selected for its representativeness:

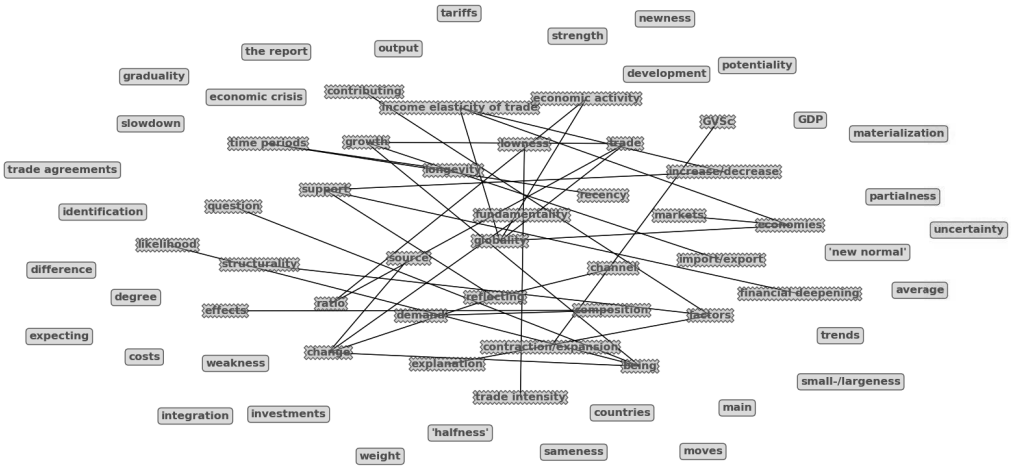

Figure 2: A diagram of chain interactions in Text E

The figure shows the web formed by the altogether 32 CISs in Text E. The majority of CISs are connected to at least one, and, in certain cases, several other CISs, which makes for a highly connected network interrupted by only six 'gaps'. Yet, the diagram also illustrates that a relatively high number of chains (almost half) do not enter into CISs and thus stand in isolation

from the 'mesh' of interacting chains. This will be further commented on in the quantitative section below.

5.5 Quantitative results for the corpus at large

Table 1 below represents the totality of quantitative findings for the corpus as a whole:

Table 1: Quantitative results

Text	1: Length (standard page of 1800 characters)	2: Total MCs (n)	3: Peripheral tokens (n / % of total MC)	4: Relevant tokens (n / % of total MC)	5: Central tokens (n / % of total MC / % of relevant tokens)	6: Non-central tokens (n / % of relevant tokens)	7: Chains (total n, CIS n, non-CIS n)
A	1.4	187	55 / 29.9 %	131 / 70.1 %	43 / 23.0 % / 32.8 %	88 / 67.2 %	26, 17, 9
B	1.8	269	84 / 27.5 %	195 / 72.5 %	64 /23.8 % / 32.8 %	131 / 67.2 %	42, 15, 27
C	2.0	277	86 / 31 %	191 / 69.0 %	92 / 33.2 % / 48.2 %	99 / 51.8 %	40, 18, 22
D	2.6	325	102 / 31.4 %	223 / 68.6 %	69 / 21.2 % / 30.9 %	154 / 69.1 %	34, 21, 13
E	3.6	520	118 / 22.7 %	402 / 77.3 %	157 /30.2 % / 39.1 %	245 / 60.9 %	67, 34, 33
F	3.8	535	99 / 18.5 %	436 / 81.5 %	186 / 34.8 % / 42.7 %	250 /57.3 %	73, 38, 35

In column 1, the length of the individual texts is indicated in terms of standard pages, one standard page being defined as 1800 characters including spaces. Column 2 provides the number of message components (MCs). In columns 3–5, the number of peripheral tokens, relevant tokens and central tokens are indicated and calculated as a percentage of the total number of message components. In column 5, this is the first of the two percentage values provided. In order to specify the ratio of central to non-central tokens, column 5 provides a second percentage representing the number of central tokens as a percentage of relevant tokens. In column 6, only one percentage value is provided which represents the number of non-central tokens as a percentage of relevant tokens. Finally, in column 7, three different values are provided: The first is the total number of chains, the second is the number of chains entering into CISs and the third is the number of chains not participating in CISs.

From the quantitative results, the following patterns emerge:

- The percentage of relevant tokens (tokens entering into chains) is relatively high for all texts, with an average of 73.1 % of total tokens. Despite variations, all texts are relatively consistent in this regard.
- The percentage of central tokens (tokens entering into CISs), however, is relatively low, with an average of only 27.7 % of total message components. This is also a consistent feature across the texts. With Text C as a partial exception, the non-central-token percentage is thus consistently high across the texts.

What is less clear is whether the numbers can be taken to reflect correlation between the length of a summary on the one hand and the relevant-peripheral ratio and the central-non-central ratio on the other. In other words, the question is whether the percentage of relevant tokens and central tokens increase with the length of the text. A scrutiny of the numbers does reveal

that the relevant-token and the central-token percentages are higher in the two longest texts (E and F) than in the two shortest texts (A and B). Yet, both texts C and D disturb this pattern: In Text C, which is slightly below average length, the central-token percentage matches that of the two longest texts (E and F), and Text D, which is just about average length, has the lowest central-token percentage of all six texts.

As for the chain statistics (column 7 in Table 1), another relatively consistent pattern emerges: In all texts, apparently regardless of length, only about half of all chains enter into CISs. In this regard, Text B is an outlier with the majority of chains not participating in CISs. Apart from this outlier, no variation, e. g. length-dependent, in the pattern is discernible. (For a discussion of the significance of the quantitative results as a whole, see Section 6).

6 Conclusion

With regard to qualitative results, the analyses have revealed that sentence-to-sentence continuity is created mainly by means of interaction between Theme-Rheme structures and Given-New structures in each sentence, and only to a limited extent by means of connectors. The analyses established that there is adequate sentence-to-sentence continuity virtually all through all texts in the corpus. As for the chain-related type of connectivity, the analyses revealed that cohesive chains are mostly lexical ones, with repetition as the main type of link in these chains. As for the structural linking of chains into CISs, it emerged that this linking mainly takes place at phrase level, either in the form of Premodifier-Head or Head-Postmodifier configurations.

From the quantitative results of the cohesion and CIS analysis, it emerged that, unlike the sentence-to-sentence continuity, the chain-related connectivity of the texts is weak. Admittedly, a high number of tokens in each text were actually seen to enter into cohesive chains, thus providing the *potential* for a high degree of connectivity. However, only to a rather limited extent did the texts feature the interaction between chains required for adequate connectivity to be present. Hasan, in fact, posits that central tokens must constitute 50 % of all tokens "for a text to be unquestionably coherent" (1984: 218), which is a threshold value that all texts in the corpus fall far short of. What the results reflect is that the texts are indeed thematically focused, i. e. to a relatively large extent organized around a number of sub-themes, linguistically manifested in (separate) cohesive chains. However, what emerges as the central problem with these texts is a lack of proper *integration* of sub-themes, as reflected in the low number of central tokens. To put this more informally – and to quote Hasan (1989: 94) again – the texts do not "[say] similar kinds of things about similar phenomena" to an extent that makes the texts sufficiently connected. We thus conclude that the NTSs do not achieve a satisfactory degree of coherence.

A possible explanation – though one that must largely remain speculation – for the problem of insufficient thematic integration resides in the very nature of summaries, i. e. in the fact that these are texts where much more extensive content and a large number of sub-themes are condensed into a very short space. Consequently, the cohesive chains are in most cases relatively short, numbering only 5.5 tokens on average.

By comparison, in a seven-page Patient Information Leaflet investigated in Hill-Madsen (2019), chains averaged 12.2 tokens, and the central-token percentage was 74.3 % of total tokens, i. e. significantly above Hasan's (1984) threshold value. Whether a direct causal connection can be assumed between chain length and central-token percentage is unlikely, of course.

Yet, what makes some degree of causality plausible is that the longer the chains, the higher the *potential* for repeated structural connections between chain members. If this hypothesis is valid, it may specifically be the relative shortness of the chains in the present corpus that militates against a sufficiently high degree of chain interconnectedness via CISs. Nevertheless, what the mapping of CIS integration shows (via charts of the network of interlinked CISs in each text, as in Figure 2 above) is that all texts feature a relatively connected, though not completely uninterrupted, network of core CISs. Thus, all the texts *are* in fact centred around a small number of relatively integrated sub-themes that coincide with the central concerns of the source text. The problem is that the thematic integration does not extend beyond that.

References

Askehave, Inger / Zethsen, Karen K. (2000): "Medical Texts Made Simple – Dream or Reality?" *Hermes – Journal of Language and Communication in Business* 25: 53–74.

de Beaugrande, Robert / Dressler, Wolfgang (1981/2016): *Introduction to Text Linguistics*. London / New York: Routledge.

Büring, Daniel (2007): "Intonation, Semantics and Information Structure." *The Oxford Handbook of Linguistic Interfaces*. Ed. Gillian Ramchand / Charles Reiss. Oxford: Oxford University Press. 445–474.

Chafe, Wallace L. (1976): "Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View." *Subject and Topic*. Ed. Charles N. Li. New York: Associated Press. 25–55.

Charolles, Michel (1997*): L'Encadrement du Discours: Univers, Champs, Domaines et Espaces*. (Cahier de Recherche Linguistique 6). Nancy: Université de Nancy 2.

Dam-Jensen, Helle / Heine, Carmen (2013): "Writing and Translation Process Research – Bridging the Gap." *Journal of Writing Research* 5.1: 89–101.

van Dijk, Teun / Kintsch, Walter (1978): "Cognitive Psychology and Discourse: Recalling and Summarizing Stories." *Current Trends in Text Linguistics*. Ed. Teun van Dijk. Berlin / New York: De Gruyter.

Firbas, Jan (1974): "Some Aspects of the Czechoslovak Approach to Problems of Functional Sentence Perspective." *Papers on Functional Sentence Perspective*. Ed. František Daneš. Prague: Academia. 11–37.

Fries, Peter H. (1983): "On the Status of Theme in English: Arguments from Discourse." *Micro and Macro Connexity of Texts*. Eds. János S. Petöfi / Emel Sözer. Hamburg: Buske. 116–152.

Givón, Talmy (1995): "Coherence in Text vs. Coherence in Mind." *Coherence in Spontaneous Text*. Eds. Morton A. Gernsbacher / Talmy Givón. Amsterdam: Benjamins. 59–116.

Greimas, Algirdas Julien (1966): *Sémantique Structurale: Recherches de Méthode*. Paris: Larousse.

Greimas, Algirdas Julien / Courtés, Joseph (1979): *Sémiotique: Dictionnaire Raisonné de la Théorie du Langage*. Paris: Hachette Université.

Gundel, Jeanette / Fretheim, Thorstein (2004): "Topic and Focus." *The Handbook of Pragmatics*. Ed. Laurence R. Horn / Gergory Ward. Malden/Oxford/Carlon: Blackwell Publishing. 175–196.

Halliday, Michael A. K. (1967): "Notes on Transitivity and Theme in English: Part 2." *Journal of Linguistics* 3.2: 199–244.

Halliday, Michael A. K. (1987/2002): "Spoken and Written Modes of Meaning." *On Grammar*. *The Collected Works of M. A. K. Halliday*. Vol. 1. Ed. Jonathan Webster. London: Continuum. 323–351.

Halliday, Michael A. K. (1989): *Spoken and Written Language*. 2nd ed. Oxford: Oxford University Press.

Halliday, Michael A. K. / Hasan, Ruqaiya (1976): *Cohesion in English*. London: Longman.

Halliday, Michael A. K. / Martin, James R. (1993): *Writing Science: Literacy and Discursive Power.* London: The Falmer Press.

Halliday, Michael A. K. / Matthiessen, Christian M. I. M. (1999): *Construing Experience through Meaning: a Language-based Approach to Cognition*. London: Cassell.

Halliday, Michael A. K. / Matthiessen, Christian M. I. M. (2014): *Halliday's Introduction to Functional Grammar*. London / New York: Routledge.

Harweg, Roland (1968): *Pronomina und Textkonstitution*. Munich: Fink.

Hasan, Ruqaiya (1984): "Coherence and Cohesive Harmony." *Understanding Reading Comprehension: Cognition, Language and the Structure of Prose*. Ed. James Flood. Newark, DE: International Reading Association. 181–219.

Hasan, Ruqaiya (1989): "The Texture of a Text." *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective.* Eds. Michael A. K. Halliday / Ruqaiya Hasan. Oxford: Oxford University Press. 70–96.

Helder, Bodil (2011/2015): *Textual Analysis. An Approach to Analysing Professional Texts*. Frederiksberg: Samfundslitteratur.

Hidi, Suzanne / Anderson, Valerie (1986): "Producing Written Summaries: Task Demands, Cognitive Operations, and Implications for Instruction." *Review of Educational Research* 56.4: 473–493.

Hill-Madsen, Aage (2019): "Coherence in Hybrid Texts: The Case of the Patient Information Leaflet Genre." *Hermes – Journal of Language and Communication in Business* 59: 153–187.

Hill-Madsen, Aage / Pilegaard, Morten (2019): "Variable Scope for Popularization of Specialized Terminology: The Case of Medico-pharmaceutical Terms." *Fachsprache – Journal of Professional and Scientific Communication* 41.1–2: 22–40.

Institute of Environmental management & Assessment (2012): "What is a Non-technical Summary?" 19.5.2020 <https://www.iema.net/assets/newbuild/documents/iema_e-briefing_7_jan_12_non-technical_summaries1.pdf>.

Jakobsen, Arnt L. (2005): "Investigating Expert Translators' Processing Knowledge." *Knowledge Systems and Translation (Text, Translation, Computational Processing)*. Eds. Helle V. Dam / Jan Engberg / Heidrun Gerzymisch-Arbogast. Berlin: Mouton de Gruyter. 173–189.

Kleiber, George (1999): "Associative Anaphora and Part-whole Relationships: The Condition of Alienation and the Principle of Ontological Congruence." *Journal of Pragmatics* 31.3: 339–362.

Krifka, Manfred (2008): "Basic Notions of Information Structure." *Acta Linguistica Hungarica* 55: 243–276.

Lambrecht, Knud (1994): *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Leonetti, Manuel Jungl (1990): *El Artículo y la Referencia*. Madrid: Taurus Universitaria.

Lyons, Christopher (1999): *Definiteness*. Cambridge: Cambridge University Press.

Mani, Inderjeet (2006): "Summarization of Text, Automatic." *Encyclopedia of Language & Linguistics*. Ed. Ken Brown. Oxford: Elsevier. 274–291.

Martin, James R. (1992): *English Text: System and Structure*. Amsterdam/Philadelphia: Benjamins.

Menzel, Katrin / Lapshinova-Koltunski, Ekaterina / Kunz, Kerstin (2017): "Cohesion and Coherence in Multilingual Contexts." *New Perspectives on Cohesion and Coherence*. Eds. Ekaterina Lapshinova-Koltunski / Katrin Menzel. Berlin: Language Science Press. 1–11.

Patton, Michael Quinn (2002): *Qualitative Methods and Research Evaluation*. 3rd ed. London / New Delhi: Sage.

Plainlanguage.gov (2011): *Federal Plain Language Guidelines.* 19.05.2020. <https://plainlanguage.gov/media/FederalPLGuidelines.pdf>.

Sanders, Ted J. M. / Sanders, Jose (2006). "Text and Text Analysis." *Encyclopedia of Language and Linguistics*. Ed. Keith Brown. Amsterdam/Boston: Elsevier. 597–607.

Sanders, Ted J. M. / Spooren, Wilbert T. M. / Noordman, Leo G. M. (1993): "Coherence Relations in a Cognitive Theory of Discourse Representation." *Cognitive Linguistics* 4: 93–133.

Taboada, María Teresa (2004): *Building Coherence and Cohesion: Task-oriented Dialogue in English and Spanish*. Philadelphia, PA: Benjamins.

Appendix 1: List of texts in the corpus

Text A:
López-Pérez, Víctor (2015): *Does Uncertainty Affect Participation in the European Central Bank's Survey of Professional Forecasters?* p. 2. 19.05.2020. <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1807.en.pdf>

Text B:
Kezbere, Linda / Maurer, Henri (2018): *Deficit-debt Adjustment (DDA) Analysis: An Analytical Tool to Assess the Consistency of Government Finance Statistics*. 3–4. 19.05.2020. <https://www.ecb.europa.eu/pub/pdf/scpsps/ecb.sp29.en.pdf>

Text C:
De Jong, Jasper / Ferdinandusse, Marien / Funda, Josip / Vetlov, Igor (2017): *The Effect of Public Investment in Europe: A Model-based Assessment*. 3–4. 19.05.2020.
<https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp2021.en.pdf>

Text D:
European Central Bank (2018): *The Use of the Eurosystem's Monetary Policy Instruments and its Monetary Policy Implementation Framework Q2 2016 – Q4 2017*. 4–5. 19.05.2020. <https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op209.en.pdf>

Text E:
European Central Bank (2016): *Understanding the Weakness in Global Trade: What is the New Normal?* 3–5. 19.05.2020. <https://www.ecb.europa.eu/pub/pdf/scpops/ecbop178.en.pdf>

Text F:
Grandia, Roel / Hänling, Petra / Lo Russo, Michelina / Åberg, Pontus (2019): *Availability of High-quality Liquid Assets and Monetary Policy Operations: An Analysis for the Euro Area*. 3–4. 19.05.2020. <https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op218~801632b377.en.pdf>

Appendix 2: Text E (full length)

[Sentence numbers inserted by us, AH-M & HD-J]

[1] In recent years, global trade has been exceptionally weak. [2] Annual import growth since 2012 has been half of what it was between 1980 and the Great Recession, and is currently recording the longest period of below-trend growth in almost half a century. [3] Notably, the same weakness has not been reflected in economic activity, which, while subdued, has not decelerated to the same extent. [4] As a result, the ratio of average imports to GDP growth – or the income elasticity of trade – has declined markedly relative to pre-crisis levels, such that the relationship of global trade and activity appears to have changed.

[5] The long period of sub-par trade growth has raised the question of whether this is a temporary deviation from trend or a longer-lasting phenomenon, reflecting fundamental structural change. [6] The question has been a prominent area of recent research and is highly relevant for central banks seeking to understand the role of external demand and international linkages

in shaping the outlook for domestic activity, potential output and inflation. [7] This paper aims to identify the main determinants of the decline in the income elasticity of trade with a view to identifying a possible "new normal" for trade growth.

[8] The change in the global income elasticity of trade between the pre-crisis period and more recent years is found to be mainly driven by two developments (see Table 1 below). [9] One source of change arises from compositional effects, such as the shift of growth in trade and economic activity towards economies with lower trade intensity, and changes in the composition of aggregate demand factors towards less tradeintensive components. [10] These shifts are not necessarily structural and could reverse in part over the medium term. [11] The other source of change relates to structural factors that are altering the fundamental relationship between trade and economic activity, such as the degree of trade liberalisation and the reliance on global value chains (GVCs). [12] These tend to be slow-moving changes reflecting fundamental shifts in the economy. [13] The main difference between these two sources is that the latter fundamentally changes the relationship between trade and economic activity at the level of individual countries or demand components, while the former changes the global income elasticity of trade by shifting the weight of activity among countries or demand components which differ in their underlying sensitivity of trade to economic activity.

[14] Compositional effects explain about half of the decline in the global income elasticity of trade. [15] The largest effect originates from the geographical composition of activity, especially the growing weight in the world economy of emerging market countries, which typically have a lower trade intensity than advanced economies. [16] This implies a weaker relationship between trade and economic activity at the global level. [17] To a lesser extent, demand composition effects have also contributed to the global trade slowdown: with import-intensive GDP components such as investment no longer growing more strongly than overall GDP, import growth has moderated. [18] The demand composition effects have been generally limited; however, as the global economy recovers, some strengthening of investment and thus the global trade elasticity might be expected.

[19] Although qualitatively less important, the second source of change in the global trade elasticity reflects structural developments, the influence of which materialises via three channels. [20] The first reflects waning support from factors that had previously contributed to global trade outpacing global output growth, including lower transportation costs, the removal of trade barriers through lower tariffs and the growing adherence to global trade agreements through the increase in WTO membership. [21] The second, related, channel reflects the moderation in the expansion of GVCs. [22] Over recent decades the rapid integration of emerging market economies into the world economy had boosted the expansion of GVCs, but the process of fragmenting production across borders was already slowing even before the Great Recession. [23] The contraction in GVCs also reflects in part rising labour costs in key emerging markets, a better appreciation of supply risk considerations in the wake of some natural disasters and an increasing move towards onshoring of production to export markets, which is partially motivated by a rise in protectionist policies. [24] The third channel reflects diminishing marginal support from financial deepening to facilitate export capacity. [25] As some of these explanatory factors are interconnected, an assessment of the marginal contributions to the trade weakness necessarily requires a degree of judgment. [26] Moreover, the future evolution of both the structural drivers and compositional developments remains uncer-

tain, and the identified explanatory factors may not capture the decline in the trade elasticity in its entirety.

[27] This paper concludes that the recent weakness in the global trade-income relationship constitutes a "new normal" for the medium-term global trade-income elasticity, and hence trade growth. [28] Some of the structural factors that supported rapid trade expansion in the past, such as reduced transportation costs, declines in tariffs, and support from financial deepening seem to have largely run their course. [29] The expansion of GVCs has stalled, and anecdotal evidence implies that, against the background of rising protectionist measures such as local content requirements, strong renewed expansion is unlikely to materialise in the medium term. [30] Other factors are more uncertain and may contribute to some cyclical upswing. [31] For instance, the dampening effect of low investment may gradually wane as the impact of negative shocks in emerging markets and oil-exporting countries unwinds or output gaps in advanced economies gradually close. [32] New trade agreements and closer integration of countries in the southern hemisphere into the world economy could give a fresh impetus to global trade. [33] Yet, while the trade elasticity of emerging market economies may over time converge with that of advanced economies, the underlying shift in the geographical composition of global economic activity from advanced towards emerging market economies is likely to persist. [34] As such, the upside potential for trade over the medium term appears to be limited, and the new normal for global trade can be expected to look broadly similar to the "weakness" observed over recent years on average.

Appendix 3: Cohesive chains in Text E

Below, the chains identified in Text E are grouped and listed in descending order according to the number of CISs they each participate in. For each chain, the number of instantiations (tokens) is indicated, and the different lexemes or other types of items representing the chain are indicated in italics. Where nothing else is indicated, the chain is exclusively represented by a lexeme identical with the name chosen for the chain, e. g. the chain 'trade' represented by item *trade*.

Four-CIS chains:
- 'globality' (23 instantiations: *globe-*, *world*, *overall*, *local*), 'change' (16: *change*, *alter-, shift-*), 'being' (15).

Three-CIS chains:
- 'trade' (22), 'growth' (14: *growth,* one referential item (*it*), *upswing*), 'increase/decrease' (14: *decelerated, declined*, *diminishing*, *increasing*, *rise*), 'factors' (13: *factors, determinants*, *drivers*), 'income elasticity of trade' (11), 'reflecting' (10), 'support' (4).

Two-CIS chains:
- 'time periods' (16: *time, period*, *years, decades, century*), 'composition' (14), 'economic activity' (12: *activity*, *economic activity*, *domestic activity*), 'economies' (13), 'ratio' (6: *ratio*, *relationship*).

One-CIS chains:
'supply/demand' (7: *demand*, *external demand*, *aggregate demand*, *supply*), 'effects' (7: *effect, impact*), '(emerging) markets' (7: *market*, *emerging markets*), 'recency' (6: *recent*), 'import/ex-

port' (6: *import*, *export*), 'contraction/expansion' (6: *contraction*, *expansion*), 'structurality' (6: *structural*), 'explanation' (5), 'source' (5), 'lowness' (5: *low*), 'GVCs' (= *Global Value Chains*) (5), 'channel' (4), 'fundamentality' (4: *fundamental*), 'trade intensity' (4), 'longevity' (3: *long*), 'question' (3), 'contributing' (3), 'protectionist measures (2), 'financial deepening' (2), 'likelihood' (2).

Zero-CIS chains:
'weakness' (6), 'countries' (6), '"new normal"' (5), 'development' (5), 'economic crisis' (4: *the Great Recession*, *crisis*), 'degree' (5: *degree*, *extent*, *levels*), 'half' (3), 'trends' (3: *trends*, *tends to*), 'GDP' (3), 'output' (3), 'identification' (3), 'investments' (3), 'main' (3), 'partialness' (3: *in part, partially*), 'strength' (3: *strong*, *strengthening*), 'costs' (3: *transportation costs*, *labour costs*), 'slowdown' (2: *slowdown*, *slowing*), 'sameness' (2: *same*), 'average' (2), 'potentiality' (2) 'the report' (3: *paper*, *Table 1*), 'difference' (2), 'weight' (2), 'small-/largeness' (2: *small*, *large*), 'materialization' (2), 'tariffs' (2), 'trade agreements' (2), 'integration' (2), 'moves' (2), 'uncertainty' (2: *uncertain*), 'newness' (2: *new*), 'graduality' (2: *gradual*), 'expecting' (2).

Appendix 4: A complete list of CISs in Text E

Below, all CISs identified in text E are listed according to the nature of the grammatical connection (at clause or phrase level) between the chains and with the number of interactions indicated for each CIS:

<u>Subject + Verb:</u>
- 'channel' + 'reflecting' (3)
- 'change' + 'being' (2)
- 'factors' + 'contributing' (2)
- 'growth' + 'being' (2)
- 'question' + 'being' (2)

<u>Verb + Subject Complement:</u>
- 'being' + 'likelihood' (2)
- 'reflecting' + 'change' (2)
- 'reflecting' + 'support' (2)

<u>Head + Postmodifier:</u>
- 'contraction/expansion' + 'GVCs' (4)
- 'increase/decrease' + 'income elasticity of trade' (3)
- 'source' + 'change' (3)
- 'income elasticity of trade' + 'economies' (2)
- 'ratio' + 'trade' (2)
- 'ratio' + 'economic activity' (2)
- 'support' + 'financial deepening' (2)

<u>Premodifier + Head:</u>
- 'globality' + 'trade' (8)
- 'globality' + 'income elasticity of trade' (5)
- 'demand' + 'composition' (4)
- 'composition' + 'effect' (4)
- 'globality' + 'economies' (4)

- 'trade' + 'growth' (4)
- 'markets' + 'economies' (3)
- 'recency' + 'time periods' (3)
- 'structurality' + 'factors' (3)
- 'explanation' + 'factors' (2)
- 'fundamentality' + 'change' (2)
- 'globality' + 'economic activity' (2)
- 'import/export' + 'growth' (2)
- 'increase/decrease' + 'protectionist measures' (2)
- 'increase/decrease' + 'support' (2)
- 'longevity' + 'time periods' (2)
- 'lowness' + 'trade intensity' (2)

Dr. Aage Hill-Madsen
Department of Culture and Learning
Aalborg University
Kroghstræde 3
DK-9220 Aalborg E
hill@ikl.aau.dk

Dr. Helle Dam-Jensen
Department of Communication and Culture
Aarhus University
Jens Chr. Skous Vej 4
DK-8000 Aarhus C
hed@cc.au.dk