



UDC: 338.23:336.74

DOI: 10.2478/jcbtp-2024-0010

Journal of Central Banking Theory and Practice, 2024, 1, pp. 223-243

Received: 28 December 2022; accepted: 31 March 2023

Piotr Szczerba *, **Anna Wojtyniak** **,
Joanna Niedźwiedzińska ***,
Wojciech Bogdanowicz ****

* National Bank of Poland,
Warszawa, Poland

E-mail:
piotr.szczerba@nbp.pl

** National Bank of Poland,
Warszawa, Poland

E-mail:
anna.wojtyniak2@gmail.com

*** National Bank of Poland,
Warszawa, Poland

E-mail:
joanna.niedzwiedzinska@nbp.pl

**** National Bank of Poland,
Warszawa, Poland

E-mail:
wojciech.bogdanowicz@nbp.pl

Monetary Policy Press Releases of 24 Inflation Targeting Central Banks – A Comparison of their Key Features and Complexity

Abstract: This paper describes the process of constructing a dataset of available online press releases related to monetary policy decisions published by central banks that have been pursuing a form of inflation targeting. A total of 3,807 documents were considered, as the review encompasses 24 central banks over the past 30 years. To prepare the dataset for text-mining analysis, a cleaning procedure has been performed, which is also presented in detail. The next step was to look at the readability of the documents in order to detect any patterns in its changes. We find that while there is a significant variation between central banks, there is no clear time trend in the readability of the monetary policy press releases, i.e. some central banks tend to have easier-to-read press releases than others, and this does not change with time. In turn, there is a weak indication that following the Global Financial Crisis outbreak and the European Sovereign Debt Crisis apogee, as well as at the time of withdrawing unconventional measures introduced in response to those two emergencies, press releases of advanced economy inflation targeters got temporarily harder-to-read. Overall, this paper can be viewed as a first attempt towards assessing qualitatively central banks' transparency, with respect to a flagship communication tool of inflation targeters, namely their monetary policy press releases.

Key words: Central bank communication, Press releases, Text mining.

JEL Code: E52, E55.

1. Introduction and research hypothesis

Central bank communication is considered to be a key ingredient of monetary policy framework – at least among inflation targeters, which is a group that we focus on. Among various communication instruments, press releases related to monetary policy decisions are, by large, the most widely used communication tool. Looking back at their beginnings, their substance was rather modest and often included just the information on the decisions taken, while currently a typical content also encompasses a more or less elaborated reasoning behind the decisions. At the same time, it is unclear whether there is any far-reaching standardisation of monetary policy press releases, especially if one wants to consider a relatively broad set of central banks. We decided to investigate this issue in more detail and over quite a long time span, applying text-mining techniques.

This paper's goal is twofold. Firstly, we describe the creation of the dataset of central banks' press releases, which is intended for further analysis using various text-mining methods. Secondly, we perform the assessment of readability, i.e. a feature of the text related to the ability of readers to understand it.

A document characterised by a higher readability will be understood by a broader public. If monetary authorities want to have a wide reach, readability would be a welcomed attribute of the press releases, especially since they constitute their flagship communication instrument.

We employ popular and relatively simple readability indices that measure text complexity mostly by comparing its quantitative features such as word and sentence length. This paper compares the readability of press releases between 24 central banks over the time span of 30 years, in order to detect any patterns in its changes. The motivation behind assessing the readability stems, among other things, from the fact that some central banks have been recently striving to simplify their language to reach a wider public. In this context, it would be interesting to see what is the gap between central banks with regard to readability and who has the most to gain.

We expect to find a marked variation between monetary authorities with some central banks having consistently easier-to-read press releases than others. Concerning evolution over time, we expect some convergence of inflation targeters with less-readable press releases towards those with more-readable press releases.

Overall, this paper can be viewed as a first step towards a qualitative assessment of central banks' transparency. This attempt is undertaken with respect to mon-

etary policy press releases which remain the key policy document and hence have the longest history of regular publication by a relatively large group of central banks. That is why our findings should be of relevance for possibly different audiences.

2. Overview of existing literature

As the weight of central bank communication has grown, so did an interest of researchers in this topic. In recent years, their motivation has been reinforced by the emergence of automated tools, in particular, text-mining methods. Those methods enable to extract and quantify certain features of various texts, including their sentiment or topics without the need to read the investigated documents. Bholat, Hansen, Santos and Schonhardt-Bailey (2015) provide a comprehensive overview of text-mining techniques used for the analysis of central bank communication.

Despite a relatively large literature embedded in the application of automated linguistic tools, very few papers have provided a comparative study across several central banks. There are, however, some exceptions that should be mentioned, e.g. Kahveci and Odabas (2016) studied changes in the tone of press releases of the US Federal Reserve (the Fed), the European Central Bank (the ECB) and the Central Bank of Turkey before and after the Global Financial Crisis, and found no major change in communication language for the latter two, whereas in the case of the Fed there was a steady decline in optimism accompanied by an increase in certainty of the language (as opposed to a more nuanced and ambiguous message). In turn, Fraccaroli, Giovannini and Jamet (2017) analysed topics and sentiments in parliamentary hearings of the Bank of England, the Fed and the ECB to see that while the central bank's objectives determined the subject of the hearing, elevated unemployment rate was associated with less focus on price stability. More recently, Armelius, Bertsch, Hull and Zhang (2019) evaluated speeches of 23 central banks and concluded that there were sentiment spillovers from central banks' communication that could not be explained by trade linkages, with the Fed having a unique influence, even among prominent banks. And Gonzalez and Tadla (2022) developed a Sentiment Score index based on press releases of 18 monetary authorities to find that for majority of them such index provides additional information which enhances prediction of future policy rates (noteworthy, in contrast to the above-mentioned papers, the sample excluded the Fed and the ECB).

Quantification of various text features provides some interesting insights into patterns in central bank communication. It is also a good starting point for a more in-depth policy analysis. In this respect, an important strand of literature focuses on the link between central bank communication, especially its tone, and the predictability of monetary policy (Hendry and Madeley, 2010). One common approach is to apply the dictionary technique which measures the frequency of using pre-defined words or phrases, which are associated with either positive or negative sentiment of the text. The drawback of this method is that readily available dictionaries, such as the dictionary by Loughran and McDonald (2011), may misclassify words in the monetary policy context. This is why some authors decide to construct their own dictionaries that better suit the specificity of central banks' communication, in particular by enabling to distinguish between the "hawkish" or "dovish" sentiment. For example, Apel and Grimaldi (2012) manually classified subject specific pairs of words and later used an automated search-and-count-words approach to generate sentiment indicator for the Swedish Riksbank's minutes. In turn, Gonzalez and Tadde (2022) utilised the Latent Dirichlet Method to build a custom dictionary for each of the 18 reviewed central banks and used a semi-automated content analysis to measure the tone of the press releases. Both papers find that sentiment indicators, which, to some extent, gauge the stance of monetary policy, provide useful additional information for predicting policy rate movements. On top of that, Nardelli, Martens and Tobback (2017) applied a unique approach by using semantic orientation and support vector machines to indirectly measure the sentiment of the ECB communication as perceived by the media, which they found to be highly correlated with the LIBOR rate.

Another strand of literature investigates the impact of communication on financial markets. In particular, Coenen et al. (2017) concluded that market uncertainty was reduced when asset purchase programme announcements by the ECB included information on the scale of the purchases, otherwise uncertainty was amplified. Moreover, state-contingent and long-term forward guidance was found to be more effective in reducing uncertainty than its open-ended or short-term alternative. Apart from that, Schmeling and Wagner (2015) showed that the tone of ECB communication matters for asset prices through the risk channel, with positive tone changes associated with an increase in bond yields and negative tone changes resulting in higher equity volatility. In turn, Stegmann (2019) who analysed three types of Fed documents (the FOMC statements, minutes and meeting transcripts) found that their sentiment did not affect treasury yields, however, several topics appearing in those documents (e.g. improving economic conditions, easing of monetary policy) affected treasury yields around the release of the FOMC statements and minutes. Lehtimäki and Palmu (2022) find that the

impact of central bank representatives' communication varies depending on depending on the role or position of the speaker.

Finally, a noticeable body of work concentrates on the cohesion of central bank communication. In this context, Ernst and Merole (2018) showed that communication of the ECB, the Fed and the Reserve Bank of Australia with regard to inflation and economic activity was consistent with their monetary policy decisions. Similar results on the ECB communication were also obtained by Jansen and de Haan (2010).

All those examples show that central bank communication may be analysed by looking at it from various perspectives. At the same time, evaluating it across a broader sample of monetary authorities is rather rare. Against this background, this paper contributes to the existing literature mainly through its comprehensive and robust dataset, comprising press releases of 24 central banks from both advanced and emerging market economies, including the most prominent central banks, such as the Fed and the ECB.

Importantly, this paper discusses the preparation of data which is key for the implementation of any text-mining methods, since analysing an unprocessed text may yield results which are affected by its technical or formatting properties. We found that in majority of existing literature this aspect is usually not discussed in-depth with very few exceptions (Shapiro and Wilson, 2019).

Next, we provide the main descriptive statistics which can be compared not only across time but also between the central banks. We compute six indices which assess the readability of press releases, similarly to Binette and Tchegotarev (2019), whose work, however, focuses solely on the Bank of Canada's Monetary Policy Report.

3. Scope of the investigated press releases

We collected press releases available online from 24 central banks which have been pursuing a form of inflation targeting¹. We only considered documents in English as documents written in other languages are more difficult to analyse

¹ Our sample includes: Australia (AU), Brazil (BR), Canada (CA), Switzerland (CH), Chile (CL), Czechia (CZ), euro area (EA), Hungary (HU), Indonesia (ID), Israel (IL), Iceland (IS), Japan (JP), Korea (KR), Norway (NO), New Zealand (NZ), Peru (PE), the Philippines (PH), Poland (PL), Romania (RO), Sweden (SE), Thailand (TH), Turkey (TR), the United Kingdom (UK), the United States (US).

and unsuitable for comparison between countries using text-mining methods. In the case of several central banks issuing two kinds of press releases on the day of the decision: shorter (with the decision) and longer (including its justification), the longer ones are taken into account². Overall, we downloaded all available English versions of press releases published until June 2020 for 24 countries which resulted in a dataset comprising of 3,932 documents.

We investigated only those monetary policy press releases that were published on a regular basis; special press releases were excluded. Among the latter were: announcements of technical details on special measures adopted to cope with crises and announcements of changes in monetary policy targets after the strategy reviews undertaken by some central banks in recent years. The reason for such an approach was that, in our assessment, those documents were not meant to belong to the same communication channel. Since we wanted to analyse press releases related explicitly to the course of monetary policy, such a decision seemed justified.

In several central banks, in the early period covered in the analysis, there were more than 12 decision-making meetings held annually, which meant that in certain months there were two regular press releases. In these cases, in order to preserve monthly frequency of the data, we calculated a mean of each statistics for such a month with two published documents being considered. As a result, the number of monthly observations was reduced to 3,807. The decision to stick to the monthly frequency was driven mainly by the motivation to use the same dataset for further more advanced research.

Due to the fact that the stating date, i.e. the date of the oldest document for each country varies, comparisons are made also for a specific sub-period of 2016-2019. The key rationale for this selection is that in this period press releases for all the analysed central banks were available³. Importantly, in 2016, all central banks have been publishing press releases following each decision-making meeting, i.e. regardless if a change in interest rates was made or not. This reinforces our argument for treating the indicated years as the most proper ones for any in-depth comparisons.

Concerning the first half of 2020, we do not include this period in the sample reviewed, because at that time many extraordinary meetings were held and the schedules of decision-making meetings were often altered at a very short notice.

² A notable example here is the ECB.

³ In this period there is no case of more than one regular press release per month.

This further restricts our sample to 3,713 observations. Needless to say, those were very busy days for central banks, with many grave decisions announced, therefore, this period requires a separate analysis that would be specifically dedicated to it.

4. Research approach

4.1. Collecting and preparing the data

To begin with, we downloaded all available English versions of monetary policy press releases from the websites of selected central banks. In most cases, documents were available as pdf or doc/docx files, sometimes as rtf files, and rather rarely as a text on a website which required pasting them into docx files.

Further analysis was performed with the use of R statistical package. After uploading the documents from individual files into R, we cleaned them to extract only their main text (mostly using the stringr package). The main text, according to our understanding, is the fundamental part of the document which provides essential information about the decision and economic rationale behind it.

In the process of cleaning the documents, all of the following parts were removed: headline, possible subheadings, postal address, website address, publication date, page numbers, lists of attendees (including any additional organisational information, e.g. on their positions), voting records, announcements about the next meeting date or upcoming publications etc., data arranged in tables or charts.

The goal of such a procedure was twofold: firstly, to preserve only the economic description relevant for explaining monetary policy; and secondly, to remove any text that could not be considered as a regular sentence, since readability indices are suited for analysing texts written in sentences, and in order to allow for a fair comparison we needed to adjust for that.

For the same reasons, if a text was presented in bullet points, we generally transformed them into regular sentences by removing bullet marks or numeration, and replacing any commas or semicolons at the end with a point . In some cases, when bullets were not written in sentences and were very short, we removed them altogether, as transforming them into sentences would require substantial interference that would affect the document to a too large extent. An example of bullets that were removed after applying this logic, was the level of several policy rates at the beginning of press releases in the National Bank of Poland.

Given a large number of press releases, changes in their layout, different file types and other irregularities, it is not possible to say with certainty that the described process is flawless. Thus, we manually compared a sample of the processed texts with the original documents and – whenever we detected more complex cases – we adjusted the cleaning procedure for a particular set of documents. Outliers were also checked separately. Still, certain omissions of parts to be deleted are possible, however, they would rather concern nonstandard one-off changes in press releases and should not have a heavy impact on the overall results.

4.2. Calculating the main text statistics

As the starting point, we computed the simplest descriptive statistics of the investigated documents, such as the number of words, sentences and paragraphs. They all relate to the main text of the monetary policy press releases, and allow for the most straightforward comparison – both across time and between central banks.

Next, text readability indices were calculated using the *quanteda* package (Benoit et al., 2018). For each document, we obtained a few most popular readability indices (see Annex for formulas): the Automated Readability Index (ARI), the Coleman-Liau Grade Level (CL), the New Dale-Chall Readability Formula (DC), the Flesch-Kincaid Grade Level (FK), the Gunning's Fog Index (GF), and Simple Measure of Gobbledygook (SMOG). For the sake of clarity, we treat the FK as the main index and use the remaining measures for robustness checks. Moreover, due to the very high correlation between the GF and the SMOG (>0.99), in what follows we report only the latter.

The construction of all of these indices relies mostly on features that can be quantitatively assessed such as the length of words and sentences. Nevertheless, the interpretation of the indices varies. Some of them are calibrated to reflect the number of years of schooling (in the US system) needed to understand a given text. For example, a Flesch-Kincaid index of 12 indicates that the text should be understood by a person with at least 12 years of schooling. Indices ARI, CL, and GF are also constructed to reflect the years of schooling, hence their interpretation is the same. In turn, the Dale-Chall formula differs from other analysed indices in two important ways. Firstly, it is the only index whose higher values are associated with a text that is easier to understand. Secondly, apart from purely quantitative components, it also analyses the share of “familiar words” from a pre-specified list, which constitutes a proxy for the qualitative features of a text. The higher the share of non-familiar words (i.e. words not included in the list) the more difficult the text is.

4.3. Disclaimers on the method used

Despite many advantages of the method applied in this paper, its certain shortcomings should be accounted for. First, the fact that we chose to analyse the English versions of documents means that, in most cases, translations of the original texts are investigated. That is why the quality of translation that depends on a number of factors, including linguistic closeness of a given national language to English, also affects our results. At the same time, the English version of any press release is functionally the primary one for international audience, hence its readability matters.

Second, readability indices use relatively simple formulas which may not always adequately reflect difficulty of a text, e.g. some complex economic concepts may be difficult to understand even if they are put in phrases comprising of short words and thus – according to the indices – should be easy for the reader. Despite their limitations, the indices do measure text difficulty and they should allow for comparisons across time and countries, since those shortcomings should be common to the whole sample considered.

Third, there are various readability indices available and their choice will likely influence results, to some extent. That is why, for robustness, we decided to use a few most popular indices. Still, it should be remembered that many other indices are proposed in the literature. Many are, however, constructed in a quite similar way to those that we include and, consequently, should show high correlation with the measures used in this paper⁴.

Fourth, while we tried to cross-check the cleaning process, to the extent feasible given a large number of observations, it is still possible that some irregularities were not captured. This especially applies to one-off differences in documents, which are more likely to be overlooked. Hopefully, such cases should be spotted as outliers, for which we have also looked in the sample.

Fifth, R does not treat various types of document files (e.g. pdf, doc, docx) uniformly. We considered this in the cleaning process and, if needed, applied additional cleaning procedures to ensure that the file type did not affect the results. In particular, it turned out that pdf files are exceptionally problematic since in a non-negligible number of documents we detected extra spaces between letters of

⁴ The correlation between various indices may be negative or positive, since some are constructed in a way so that “the more difficult the text, the lower the index value”, while other are constructed in a way so that “the more difficult the text, the higher the index value”.

a word. Another issue with pdf files is that when uploaded into R headlines or footnotes of those documents may appear in the middle of the main text, as was the case e.g. for the press releases of the Bank of Iceland. All of this was accounted for to the best of our knowledge, therefore, these particularities should not have a significant bearing on the findings.

Finally, certain central banks temporarily published noticeably more complex documents which included a lot of additional information. Such instances required making decisions on how far should the cleaning procedure go (i.e. what to exclude) so that the interference is not biasing the results. While such cases inevitably led to applying more judgmental approach, we aimed at ensuring a level playing field and mostly removed additional information.

5. Results – descriptive analysis

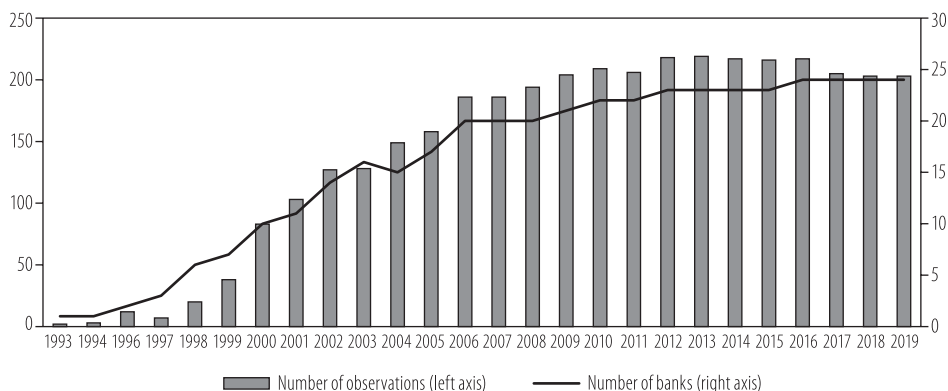
The earliest observation date varies between monetary authorities and corresponds to the oldest document that was available at the central banks' website in English. That is why in order to get a comparable sample we used a carefully selected period in which documents for all the analysed central banks were available. According to this logic, press releases published in 2016-2019 were investigated (Table 1).

Table 1: Number of documents published by central bank

	AU	BR	CA	CH	CL	CZ	EA	HU	ID	IL	IS	JP
Start of sample	1993	2006	1997	2000	1998	2010	1998	2002	2016	1998	2009	2012
Number of observations until 2019 (total)	165	111	169	87	238	81	229	201	45	230	87	80
Number of observations per year (median)	6.0	8.0	8.0	4.0	12.0	8.0	11.0	12.0	12.0	12.0	8.0	10.0
	KR	NO	NZ	PE	PH	PL	RO	SE	TH	TR	UK	US
Start of sample	1999	2001	2006	2002	2002	2003	2003	1996	2000	2006	2000	2005
Number of observations until 2019 (total)	225	147	108	213	163	181	127	157	165	156	227	121
Number of observations per year (median)	12.0	8.0	8.0	12.0	8.5	11.0	8.0	6.0	8.0	12.0	12.0	8.0

Source: Own elaboration based on central banks' documents.

Figure 1: Number of central banks publishing monetary policy press releases and number of documents issued in 1993-2019

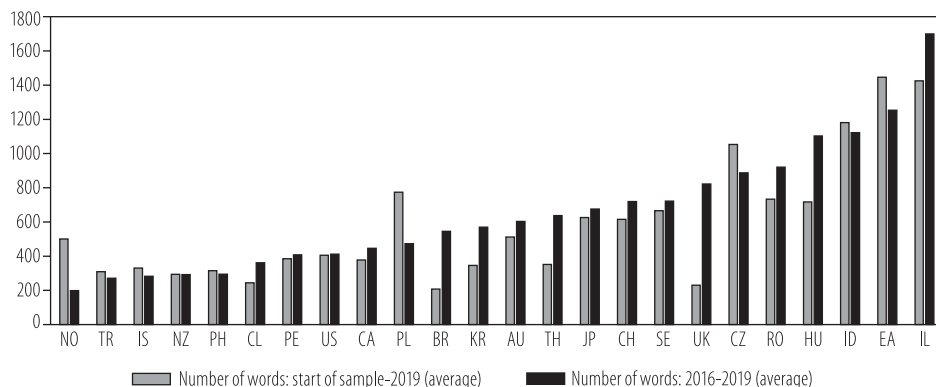


Source: Own elaboration based on central banks' documents.

The fall in the number of central banks issuing monetary policy press release recorded in 2004 (as compared to 2003) results from the Reserve Bank of Australia not publishing any press releases in 2004 (at that time press release accompanied only changes in policy rates, yet in 2004 policy rates were kept unchanged)

The median number of observations per year (Table 1) gives a rough approximation of a frequency of decision-making meetings. In this context, two reservations should be mentioned. First, our decision to stick to monthly frequency means that if a given central bank held more than one decision-making meeting in a certain month, an average of values and characteristics obtained for all the documents issued in this period is computed. Second and possibly more important, the frequency of decision-making meetings has not been constant over time. In particular, around 2010 one can observe a shift towards 8 meetings per year (from 12 preferred previously) that later spread to quite a few central banks considered in our sample (Niedźwiedzińska, 2022). Thus, by the end of the analysed period, there is a decrease in the total number of observations (Figure 1), which to great extent stems from the lower frequency of decision-making meetings.

In some cases (Australia, Canada, Chile, Romania, and Sweden), in the early years covered by the overall sample, the press releases were published only when a change in policy rates took place. This could affect the comparison, in particular with the central banks issuing press releases irrespective of the decision taken. Over time there was a move to a regular publication of monetary policy press releases, i.e. after every decision-making meeting. Hence, in the period 2016-2019, every decision-making meeting was accompanied by a dedicated press release.

Figure 2: Number of words in the main text of the documents (averages)

Source: Own elaboration based on central banks' documents.

Series referring to "start of sample-2019" embrace a different time period for each central bank, depending on the earliest available observation, so their comparability is limited.

The length of documents varies considerably between central banks (Figure 2). Some documents, especially the older ones, include only a very short explanation of the decision (or just a single sentence, e.g. the Bank of England's early press releases), while others are several paragraphs and few pages long⁵. For some monetary authorities, significant changes in the length of press releases are noted over time (this is the case of, e.g., Norway, Poland, Brazil, Korea, Thailand, the United Kingdom). These modifications go both ways – i.e. certain central banks, at some point, opted for shorter, while others opted for longer press releases than issued in the past. As a result, there is no clear pattern to be observed.

6. Results – readability indices

6.1. Comparison between central banks

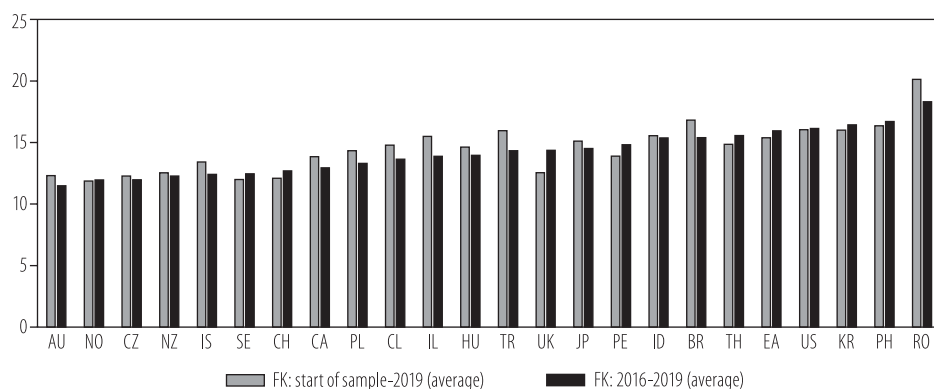
Readability indices also point to visible differences between central banks. Some tend to lead the ranking of having the most readable press releases across time (Figure 3). These include monetary authorities of Australia, Norway and Czechia.

⁵ On top of that, some press releases have annexes or tables with additional data or information. However, these are not considered a part of the main text (see section 5.1), so they have no influence on the calculated text statistics, such as length and readability indices.

At the other end of the scale are central banks of Romania, the Philippines and Korea. As shown in Figure 4, some central banks have consistently less readable press releases than others, even though in several cases they are displaying significant shifts in their readability scores over the years.

The comparison of the readability indices allows to spot some patterns and to formulate certain hypotheses concerning the lack of homogeneity. Generally, central banks in advanced economies tend to have easier-to-read press releases than central banks in emerging market economies (Figure 5), which may imply that advanced economies possibly have better access to staff more fluent in drafting, or more aware of the benefits stemming from clear communication. At the same time, the globally prominent central banks, the Fed and the ECB, are placed in the lower half when ranked by readability. Similarly, central banks of other relatively large advanced economies, Japan and the United Kingdom, perform slightly below the median. This, in turn, may suggest that monetary policy communication of the major central banks is meant to be addressed to more sophisticated observers. Thus, overall, central banks from small advanced economies score best.

Figure 3: Flesch-Kincaid Grade Level for individual central banks (averages)



Source: Own elaboration based on central banks' documents.

Series referring to "start of sample-2019" embrace a different time period for each central bank, depending on the earliest available observation, so their comparability is limited.

Figure 4: Flesch-Kincaid Grade Level for individual central banks (averages of monthly observations against volatility range)

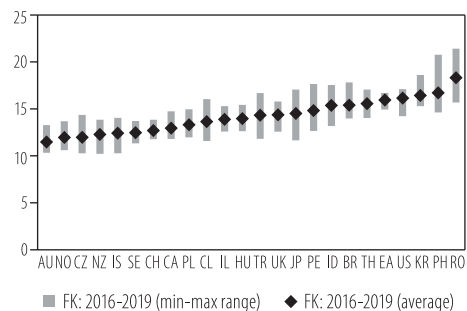
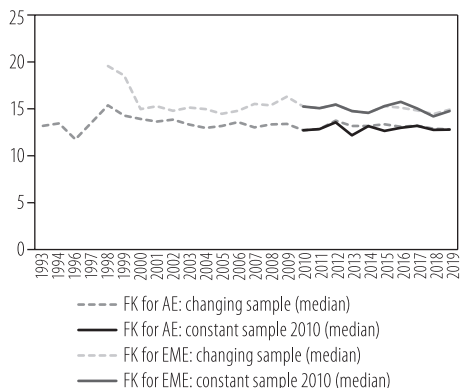


Figure 5: Developments in Flesch-Kincaid Grade Level in 1993-2019 (median of annual averages)



Source: Own elaboration based on central banks' documents.

Series referring to "constant sample 2010" includes only countries for which we have observations for 2010. Thus, in subsequent years this group is fixed, irrespective of new countries entering the overall sample (named here "changing sample"). "EME" stands for emerging market economy inflation targeters. "AE" stands for advanced economy inflation targeters.

Additionally, we checked whether text length affects its readability. As expected, there is a very low correlation between text length (measured in the number of words) and readability. In turn, somewhat surprisingly, there is no indication that central banks which use English as the primary language score better compared to central banks in other advanced economies.

6.2. Comparison across time

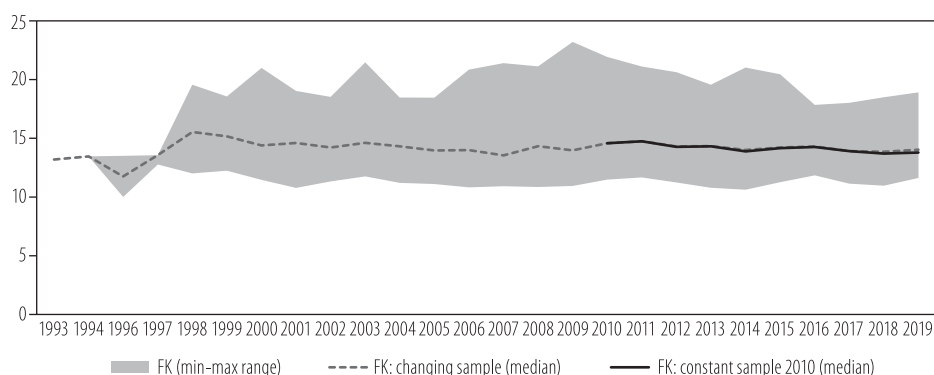
In general, the readability indices in the pooled sample show no trends. Likewise, in most cases there are no clear tendencies to be noted at the individual central bank' level. This would suggest that there is no visible move towards better readability.

Moreover, contrary to our expectations, we do not see evidence of convergence towards better readability. In most cases, central banks with less readable press releases at the start of their publications do not show a clear improvement in readability score going forward. Even though, there are some exceptions to that rule; notably, the Bank of Israel's press releases became significantly easier to

understand around 2005⁶. Examples of other central banks which have seen an improvement over the years include Turkey and Iceland. Interestingly, the Fed's press releases went in the opposite direction, as after the period of relatively good readability before 2008, they became significantly less readable⁷. Peru had also seen a deterioration in the readability for much of the analysed period, but it improved in 2018-2019.

In fact, this may be related to another observation, namely that press releases got more difficult to understand in the crises years (Figure 6). In particular, in the early months of the Global Financial Crisis (after the fall of the Lehman Brothers) and at high times of the European Sovereign Debt Crisis, as well as when unconventional measures introduced in response to those two emergencies were withdrawn, communication of many central banks, and especially of advanced economy inflation targeters, became more difficult compared to previous years⁸. However, this is not consistent across all the analysed central banks and requires further research, since on such occasions central banks often issued special press releases or held additional meetings concluded with extra press releases.

Figure 6: Developments in Flesch-Kincaid Grade Level in 1993-2019 (median of annual averages against volatility range)



Source: Own elaboration based on central banks' documents.
Shaded area indicates the range of individual banks' results.

⁶ Flesch-Kincaid Grade Level score for the Bank of Israel's press releases in the period 1998-2004 averaged 18.7, while in the period 2006-2019 it declined to 14.1.

⁷ Flesch-Kincaid Grade Level score for the Fed's press releases in the period 2005-2007 averaged 13.4, while in the period 2009-2019 it picked to 16.9.

⁸ A similar pattern could be observed for the early months of the COVID-19 pandemic, which is, however, not covered by the analysis.

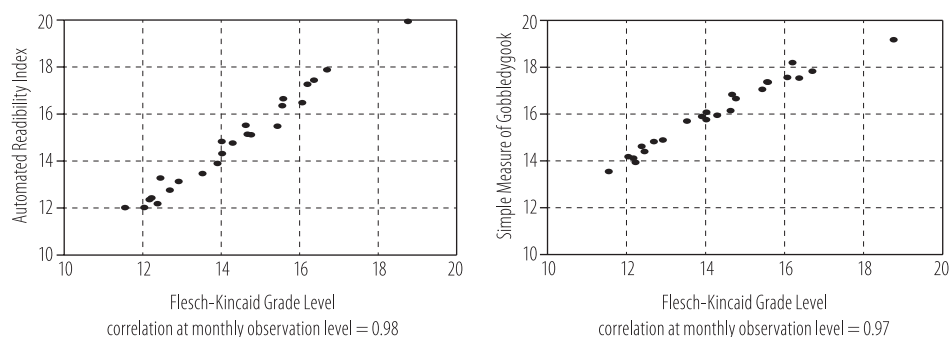
These often differ significantly in structure from regular press releases which could affect the comparison. As explained earlier, in order to avoid such disturbances we considered only regular press releases in the analysis.

7. Robustness checks

In order to check whether choosing the Flesch-Kincaid Grade Level as the key measure could affect the results, we calculated correlations between this index and four other discussed earlier. Below we show correlations between indices computed separately for each central bank for the period 2016-2019 (Figure 7).

The correlations between the Flesch-Kincaid Grade Level and both the ARI and the SMOG are very high. Likewise, the Dale-Chall Readability Formula shows high (though negative, due to the index' construction) correlation with the FK. In turn, the Coleman-Liau Grade Level is visibly less correlated with the Flesch-Kincaid index. While the latter two indices have similar construction to the Flesch-Kincaid Grade Level, they differ in details (see Annex). The Dale-Chall Formula – as already noted – is not purely quantitative. The lowest correlation with the Coleman-Liau Grade Level may follow from relatively higher weight attached to word length and lower weight attributed to sentence length, compared to other indices – a consequence of this is a much better result obtained for Romania which uses very long sentences.

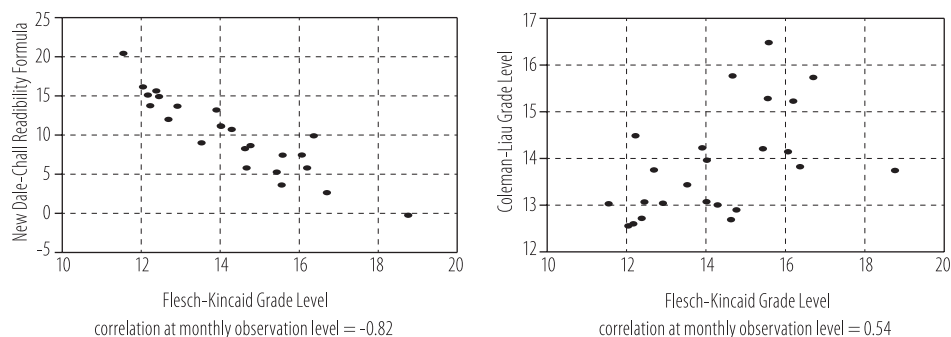
Figure 7: Comparison of Flesch-Kincaid Grade Level and other readability indices



Source: Own elaboration based on central banks' documents.

A dot represents an average readability score for a given central bank for the period 2016-2019. The Dale-Chall Readability Formula is constructed in the way that higher score means easier-to-read document.

Figure 7: Comparison of Flesch-Kincaid Grade Level and other readability indices - continued

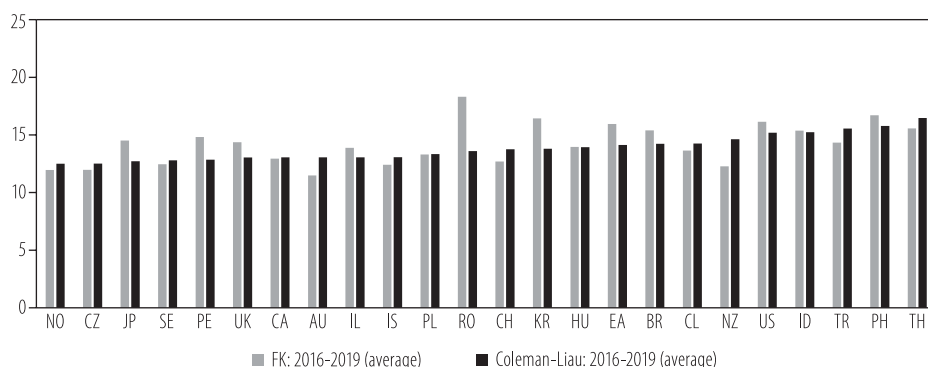


Source: Own elaboration based on central banks' documents.

A dot represents an average readability score for a given central bank for the period 2016-2019. The Dale-Chall Readability Formula is constructed in the way that higher score means easier-to-read document.

Figure 8 shows changes in the central banks' ranking based on their press releases readability when the most diverging indices are used, i.e. the Flesch-Kincaid and the Coleman-Liau Grade Levels. For many countries the choice of specific formulas does not materially affect their position in hierarchy, but for some it does. In particular, based on the Coleman-Liau scores, results for Romania, Korea, the euro area, Japan and Peru are much better and that for New Zealand is much worse compared to the Flesch-Kincaid Grade Level indications.

Figure 8: Comparison of Flesch-Kincaid Grade Level and the Coleman-Liau Grade Level for individual central banks



Source: Own elaboration based on central banks' documents.

8. Conclusions and areas for further research

We have gathered an extensive dataset of press releases from central banks that have been pursuing a form of inflation targeting. The dataset contains available online press releases in English that we transformed into a text-mining input by cleaning and standardising the documents. We strived to extract the main text which contains the description of the decision, written in full sentences, excluding additional parts such as data tables, technical and address information, etc. This paper contains in-depth description of the steps that led to the creation of the dataset, including cleaning procedures and choices that were made in the process.

This paper also provides an analysis of readability of the press releases. We mostly use the Flesch-Kincaid Grade Level scores, but we cross-checked them with results based on four other commonly applied readability indices. The comparison between press releases of the reviewed central banks shows that there is a significant variation in how clearly monetary authorities communicate using this particular instrument of conveying their message. Central banks from small advanced economies tend to have better readability, while emerging market economies – as well as large advanced economies – usually perform worse than average. Additionally, we explored also trends over time, but did not see any consistent tendencies towards easier communication for most of the central banks although there are certain exceptions. Neither did we find evidence for convergence among central banks.

The strength of the text-mining method lies in the possibility to compare large datasets which could not be done otherwise. At the same time, this method has some limitations as well and, in particular, using big datasets means that results at an individual press release level in certain cases may not be precise. However, we had undertaken a lot of effort to ensure high quality and comparability of the sample.

Therefore, the data we collected allow for further analysis, including more advanced modelling, investigation of topics and sentiments, etc. Also readability analysis could be expanded, e.g. by exploring reasons for the differences that we have found to explain why some central banks are characterised by better readability and what is the cause of readability changes over time.

References

1. Apel, M. and Grimaldi, M.B. (2012). The Information Content of Central Bank Minutes, *Sveriges Riksbank Working Paper Series*, 261.
2. Armelius, H., Bertsch, Ch., Hull, I. and Zhang, X. (2019). Spread the Word: International spillovers from central bank communication, *Journal of International Money and Finance*, 103 (2020).
3. Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S. and Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30), 774, available at: <https://doi.org/10.21105/joss.00774>.
4. Bholat, D., Hansen, S., Santos, P. and Schonhardt-Bailey, C. (2015). Text mining for central banks: handbook. *Centre for Central Banking Studies* (33): 1-19.
5. Binette, A. and Tchegotarev, D. (2019). Canada's Monetary Policy Report: If Text Could Speak, What Would It Say?, *Bank of Canada Staff Analytical Note*, 2019-5.
6. Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
7. Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E. and Strasser, G. (2017). Communication of monetary policy in unconventional times, *CFS Working Paper Series*, No. 578.
8. Ernst, E. and Merole, R. (2018). Central bank communication: A quantitative assessment, International Labour Organization, *Research Department Working Paper*, No. 33.
9. Fraccaroli A., Giovannini, A. and Jamet, J-F. (2017). Central banks in parliaments: a text analysis of the parliamentary hearings of the Bank of England, the European Central Bank and the Federal Reserve, *ECB Working Paper Series*, No 2442.
10. Gonzalez, M. and Tadla, R. C. (2022). Monetary policy press releases: an international comparison, *BIS Working Papers*, No 1023.
11. Hendry, S. and Madeley, A., (2010). Test Mining and the Information Content of Bank of Canada Communications, *Bank of Canada Working Paper*, 2010-31.
12. Jansen, D.J. and de Haan, J. (2010). An Assessment of the Consistency of ECB Communication using Wordscores, *DNB Working Paper*, No. 259.
13. Kahveci, E. and Odabas, A. (2016). Central banks' communication strategy and content analysis of monetary policy statements: The case of Fed, ECB and CBRT, *Social and Behavioral Sciences*, 235 (2016).

14. Lehtimäki, J. and Palmu, M., (2022), Who Should You Listen to in a Crisis? Differences in Communication of Central Bank Policymakers, *Journal of Central Banking Theory and Practice*, 11(3) 33-57.
15. Loughran, T. and McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks., *Journal of Finance*, 66: 35-65.
16. Niedźwiedzińska, J. (2022). Inflation Targeting and Central Banks: Institutional Set-ups and Monetary Policy Effectiveness, Routledge.
17. Nardelli, S., Martens, D., and Tobback, E. (2017). Between hawks and doves: measuring central bank communication, *ECB Working Paper Series*, No 2085.
18. Schmeling, M. and Wagner, C. (2015). Does Central Bank Tone Move Asset Prices?, available at SSRN: <http://dx.doi.org/10.2139/ssrn.2629978>.
19. Shapiro, A.H. and Wilson, D.J. (2019). Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analysis, Federal Reserve Bank of San Francisco Working Paper Series, 2019-02.
20. Stegmann, J. (2019). Federal Open Market Committee communication: a text mining analysis, Honors Theses.

Annex

Calculated indices; based on quanteda package (Benoit et al., 2018):

- Automated Readability Index

$$ARI = 0.5 \times ASL + 4.71 \times AWL - 21.34$$
- Coleman-Liau Grade Level

$$CL = -27.4004 \times CL.ECP \times 100 + 23.06395$$

$$CL.ECP = 141.8401 - 0.214590 \times 100 \times AWL + 1.079812 \times (nst \times 100) / nw$$
- New Dale-Chall Readability Formula

$$DC = 64 - (0.95 \times 100 \times (nwd / nw)) - (0.69 \times ASL)$$
- Flesch-Kincaid Grade Level

$$FK = 0.39 \times ASL + 11.8 \times (nsy / nw) - 15.59$$
- Gunning's Fog Index

$$GF = 0.4 \times (ASL + 100 \times (nwsy \geq 3 / nw))$$
- Simple Measure of Gobbledygook

$$SMOG = 1.043 \times (nwsy \geq 3) - 0.5 \times (30 / nst) + 3.1291$$

Where:

- nw = number of words
- nsy = number of syllables
- nst = number of sentences
- nwsy ≥ 3 = number of words with 3 syllables or more
- nwd = number of “difficult words” not matching the Dale-Chall list of “familiar words”
- ASL = Average Sentence Length: number of words / number of sentences
- AWL = Average Word Length: number of characters / number of words