Fraudulese:

Analysing the language of fraud

A linguistic study revealing how language can be used by fraudsters in short messages.

Linguistic research commissioned by Visa and conducted by researchers from the Aston Institute for Forensic Linguistics

April 2022

Foreword

By Mandy Lamb, Managing Director, Visa UK & Ireland

With so many platforms and online marketplaces to browse and buy from, online shopping puts a world of possibilities at our fingertips. However, as we're all spending more time online, it's good to be aware of what we can do to keep ourselves safe.

Research shows that more than a third of online shoppers (35%) have been a target of fraud, with over half (55%) reporting an increase in contact from fraudsters over the past 12 months.[1]

This is backed up by data from the UK Government which revealed that over 100,000 shoppers in the UK have fallen victim to fraud since December 2020 – a loss of £60 million.[2]

Almost all (97%) UK shoppers regularly buy items online, and as our lives become increasingly digital, fraudsters are also evolving their techniques to keep up with the latest advances in technology.[3]

At Visa, we are committed to protecting people from online payment fraud. That's why we've partnered with researchers from the Aston Institute for Forensic Linguistics to undertake an industry-first analysis of short fraudulent text messages, emails, and social media messages, to launch 'Fraudulese' – the language of fraud.

Along with Visa's own survey of 2,000 UK consumers, which explores the awareness of fraud among shoppers and the extent to which they experience it, we're working to equip consumers with the skills to recognise the tell-tale signs of fraud and stay safe when shopping online.

When it comes to paying with Visa, you can feel confident you are paying safely and securely, as Visa's Zero Liability Policy* means you won't be held responsible for unauthorised or fraudulent charges made with your account. We are constantly evolving our security and payment technology to ensure that online payments are quick, simple, and secure so that everyone can pay with confidence.

* Visa's Zero Liability policy does not apply to certain commercial card and anonymous prepaid card transactions or transactions not processed by Visa. Cardholders must use care in protecting their card and notify their issuing financial institution immediately of any unauthorized use. Contact your issuer for more detail.

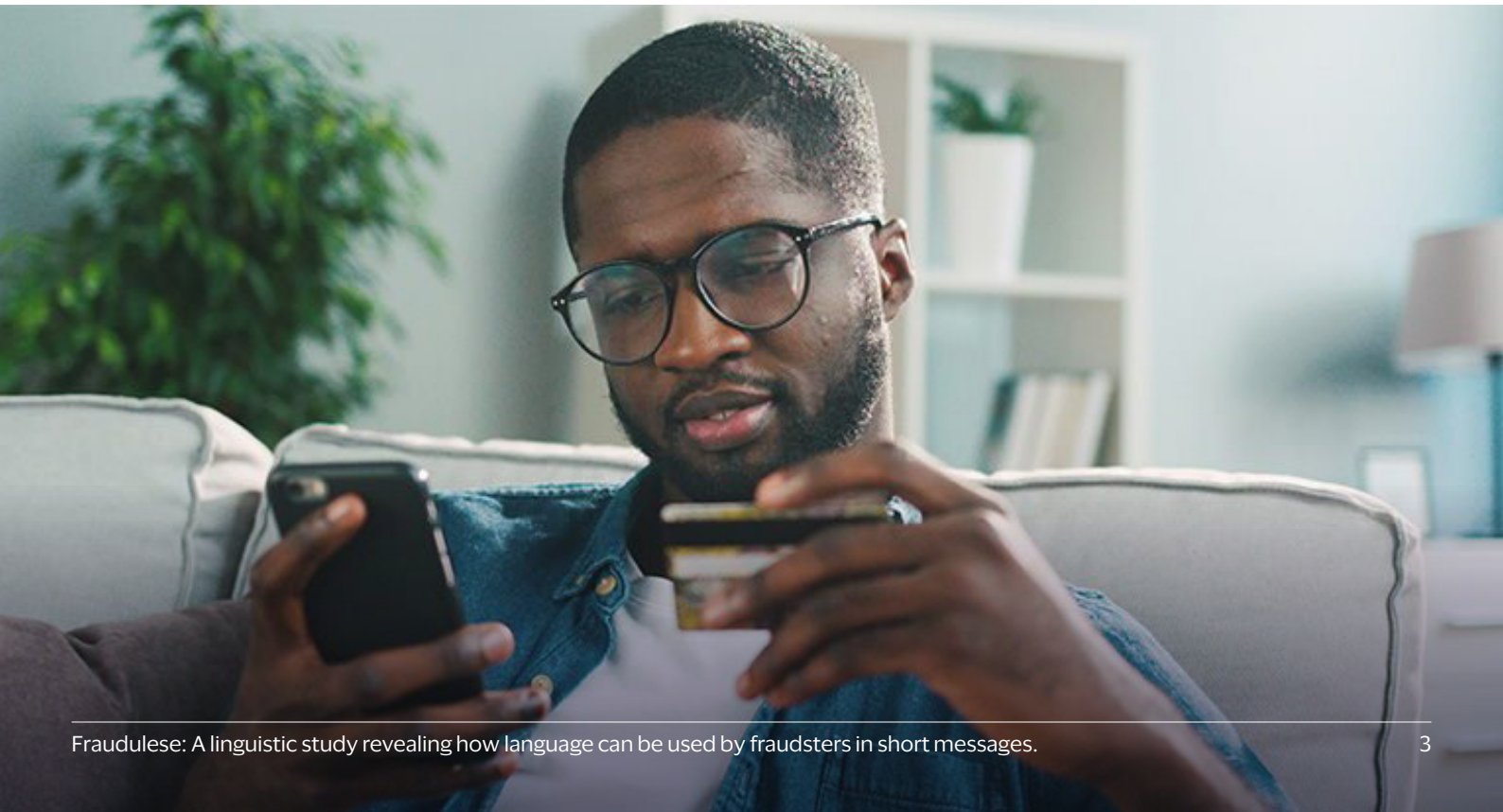
Introduction

This research report analyses a dataset of 155 fraudulent messages (5,891 words), identified as 'phishing' and 'smishing' scams in the form of email and text messages.[4]

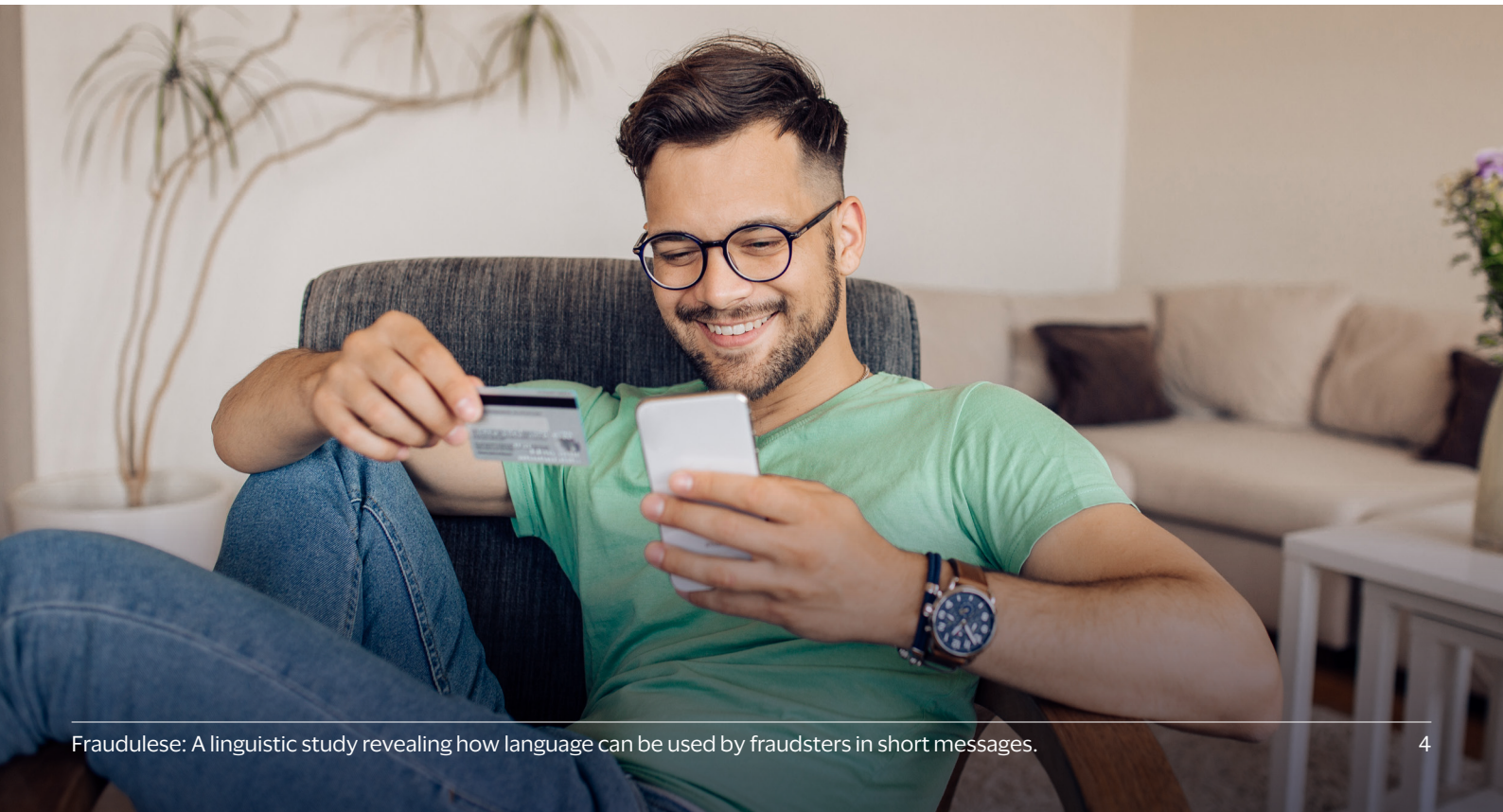
It reveals the most common words and phrases used by fraudsters to target victims, and offers advice for shoppers about how to stay vigilant. The COVID-19 pandemic has seen a particularly steep rise in fraudulent messages from people claiming to be health services and delivery companies in attempts to steal private information.[5]

Both of these scam types are present in the dataset analysed in this report, as well as messages imitating commercial companies, banks, and private individuals.

This analysis is also supported by consumer research of 2,027 nationally representative UK adults to understand more about current consumer behaviour and attitudes towards fraud.



Contents



What do we already know about the language of fraud?

Visa's consumer research finds that three quarters (72%) of British people have been targeted by what they believe to be a fraudulent message. Of these, more than half (55%) have found that the number of fraudulent messages has increased over the last year with the average UK consumer being targeted twice a week [1].

Before carrying out their study, the researchers from the Aston Institute for Forensic Linguistics reviewed existing academic literature to find out what we already know about the communicative strategies used by fraudsters. Previous academic research finds there are four most common strategies found in written fraudulent texts, listed below.

This is also supported by Visa's consumer research: when survey participants were shown a series of fraudulent messages, one in six (16%) respondents trusted the communications as legitimate, rising to one in four (25%) 18- to 35-year-olds [1]. Amongst those who would trust the fraudulent message, the main reasons behind this were finding the wording to be trustworthy (39%), feeling the action required was clear (36%) and recognising the brand name or product (34%) [1].

According to those who have received fraudulent messages, the most common platforms are emails (79%), text messages (68%) and phone calls (47%) [1].

The first-of-its-kind analysis undertaken in this report looks to build on existing research to try to uncover the linguistic features of fraudulent communications within short, one-off messages.

> I often get phishing emails trying to get me to reset my social media account, but the most alarming fraud experience I've had was health-related. I got a text one day from a mobile number pretending to be my health provider and telling me to go to a website to update my details.
>
> The website looked identical to the healthcare site, but I had a sneaky suspicion it wasn't real as texts from my health provider aren't from a mobile number. So I found the phone number of my health provider on their website and when I spoke to them they confirmed it was likely a scam."
>
> **Emily, London, in her 20s**

Previous academic research[3] highlights four strategies commonly found in written fraudulent texts, including:

- Attempts to establish senders' credibility
- Requests for confidentiality or secrecy
- Appeals for urgency
- References to trust[6]

Understanding fraudsters' language use

When receiving an email, message or text about a product or service, Visa found that half of British people (50%) would trust it is legitimate if they recognised the email address or phone number[1].

Other common features that make a message seem trustworthy include if it's a company that has been interacted with before (49%), correct use of language and grammar (44%), the message being formatted correctly (40%) and the sender being familiar (38%)[1].

With this in mind, the researchers first explored the most frequent words, phrases, and collocations in fraudulent texts using computer-assisted analysis to identify recurring themes.

Glossary

Phrases are multi-word expressions, such as thank you or one time passcode

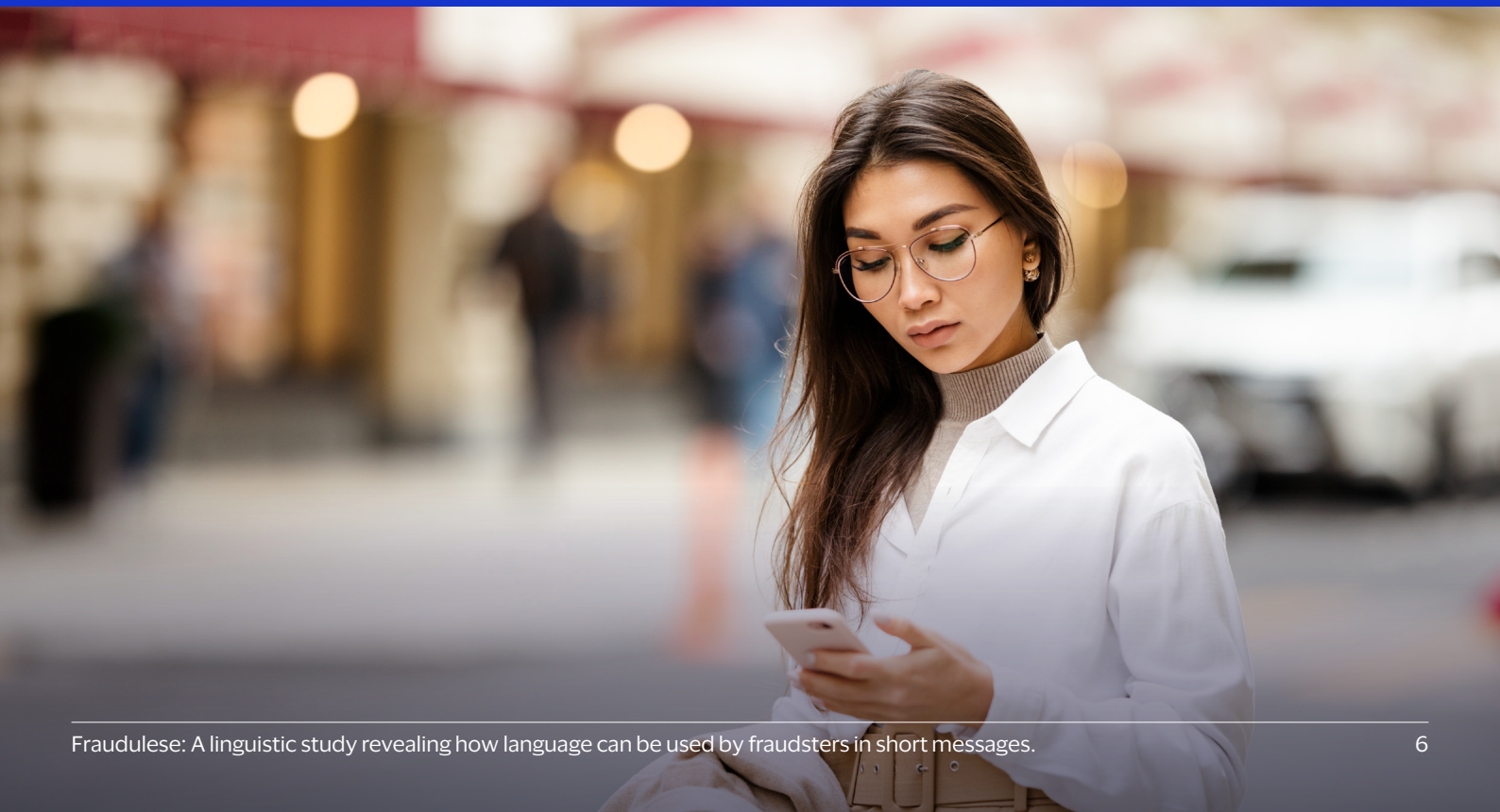
Collocations are made up of two words that frequently appear together, but not necessarily directly next to each other, such as update [...] information

Content words are words, such as nouns, verbs, adjectives and adverbs, that have a lexical meaning

The top ten content words (nouns, verbs, adjectives, adverbs) the analysis identified were:

Account
Here
Information
Email
Now
Delivery
Need
Send
Payment
Visit

See table 1 in the appendix for the top 30 content words analysed in the fraudulent messages.



The top ten phrases – combinations of two or more words – the analysis identified were:

Click here

Account information

Gift card

Email address

48 hours

Now available

Billing information

Withdrawn account

Winning prize

Can't send

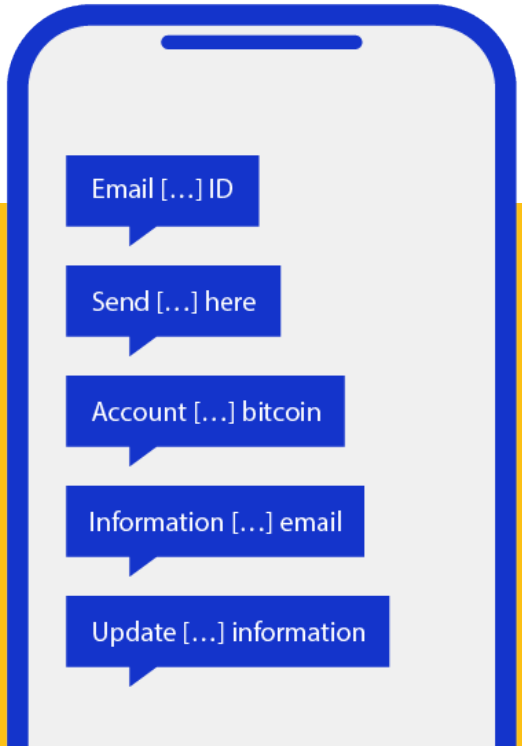
See table 2 in the appendix for the top 20 most frequent phrases analysed in the fraudulent messages.

Following the analysis of the words used in the dataset, the researchers identified six key themes in fraudulent messages:

- **Services** – for example: delivery, package, parcel, order
- **Payment** – for example: payment, card, bitcoin, pay, reward
- **Action required** – for example: need, send, visit, click, confirmation, confirm, select, claim
- **Personal information** – for example: account, information, email, detail, name, ID
- **Urgency** – for example: now, ready, hour
- **URL links** – links are more frequent than all of the words above, except 'account'. Fraudulent links tend to be lengthy and sometimes not relevant to the company the sender is claiming to be

Whilst many genuine texts related to services and payment will most likely feature these words, if a message contains multiple of the words listed above, it may be a sign of fraud.

What's more, two thirds of British people (66%)[1] say that the top giveaway of a fraudulent message is asking for money or bank details. The research backed this up by revealing the most common phrases and collocations in the analysed texts, many of which relate to account details.



The top five collocations our analysis identified were:

See table 3 in the appendix for the top 20 most frequent collocations analysed in the fraudulent messages.

Highlights

Following the identification of these top phrases and collocations, the researchers identified four key themes in fraudulent messages.

With Visa's consumer research revealing that one in five (23%)[1] consumers would only read a message in 'some' detail before following instructions or clicking on any links, it's important that consumers are equipped with the tools and knowledge to spot the potential signs of fraud.[8]

Four key themes in fraudulent messages:

Resolving a problem

Phrases include: billing information, withdrawn account, can't send, account expires, missing information, shipping.

Sharing positive news

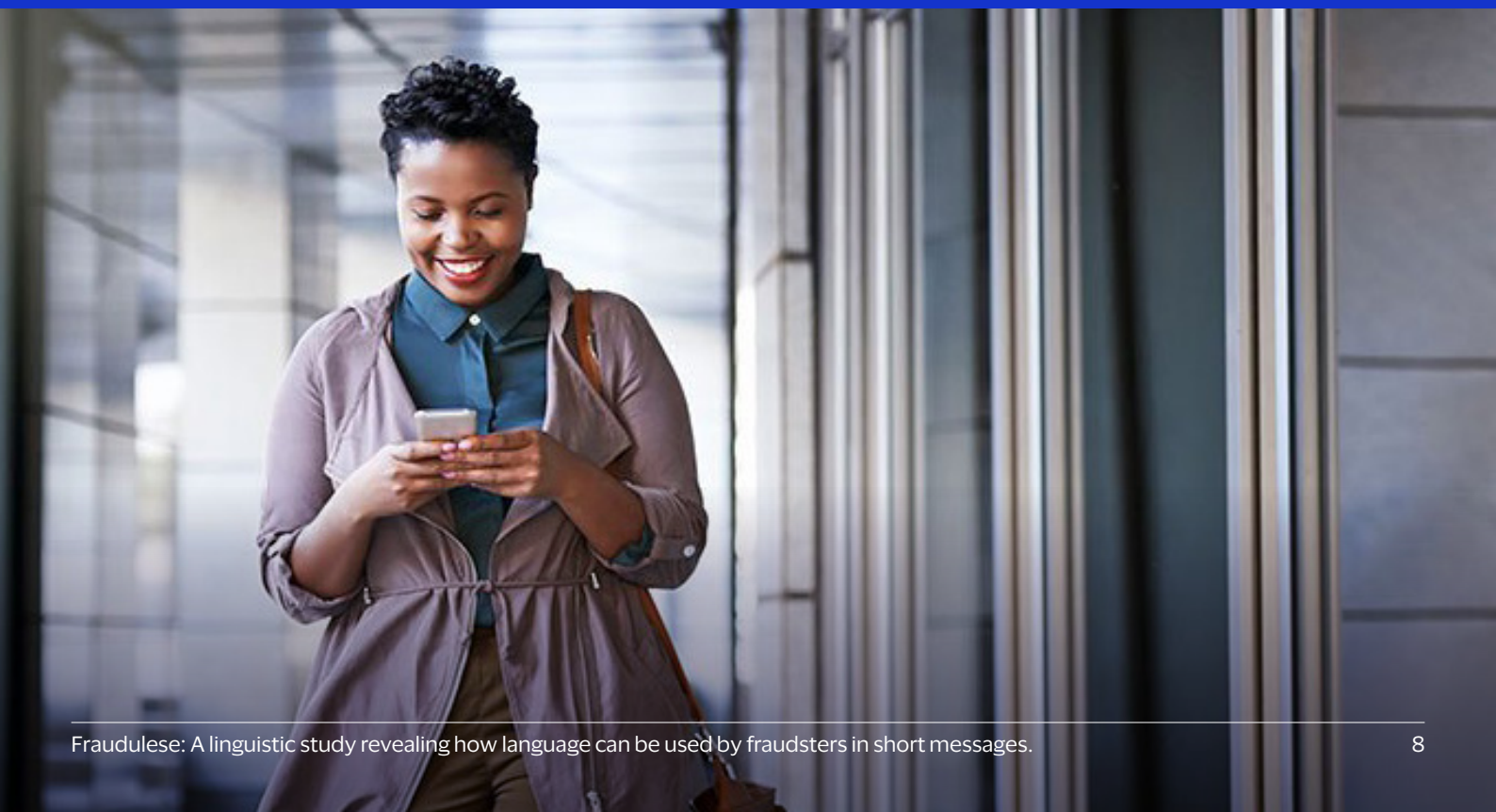
Phrases include: gift card, now available, winning prize, bitcoin program, bitcoin bonus, account bitcoin.

Taking action

Phrases include: click here, confirmation process, see profiles, send [...] here, update [...] information, click [...] below, click [...] button, verify [...] account, send [...] details, verify [...] email, confirm [...] account.

Sense of urgency

Phrases include: 48 hours, tomorrow morning.



Understanding fraudsters' communicative strategies

Visa's consumer research revealed that nine out of ten consumers (86%)[1] are confident they could recognise that a message is from a fraudster, but with over 100,000 shoppers falling victim to fraud since December 2020[9], understanding more about the language fraudsters use is important to help shoppers protect themselves.

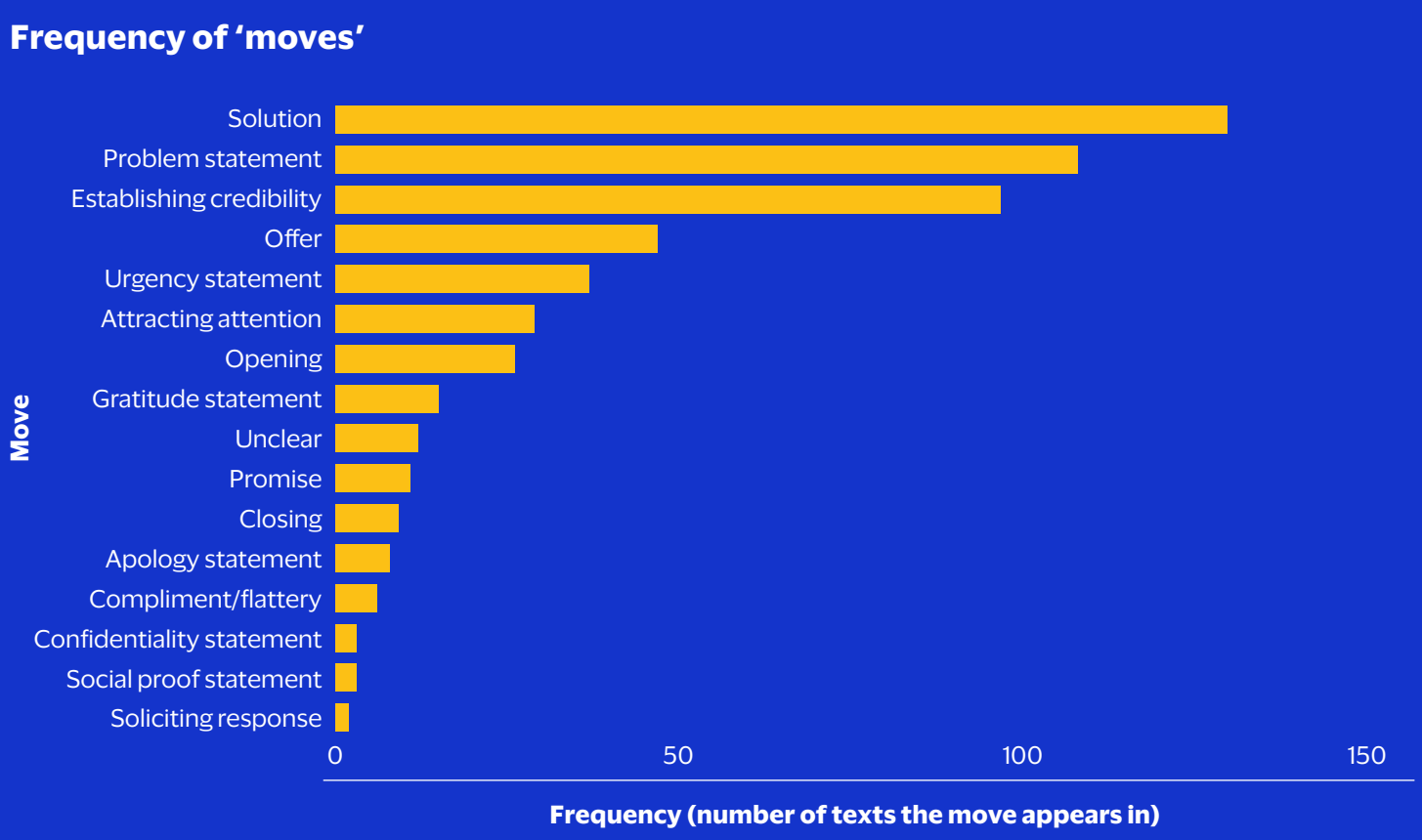
To offer a more comprehensive view of the language of fraudulent texts in the dataset, the researchers used a method known as 'move analysis'.

Move analysis involves breaking down the texts into functional units (e.g., Offer, Problem statement), to reveal the common communicative strategies used in fraudulent messages.

Glossary

Move: a unit of text with a specific function such as Solution, Offer, Problem Statement etc.

Figure 1. Frequency of 'moves' – a unit of text with a specific function



The top five most common moves in the analysed fraudulent messages are:

MOVE 1

Solutions

Solutions are the most common move, occurring in 87% of the texts. These invite you to respond to a Problem statement or Offer, either through an instruction or suggestion, and usually involve clicking a link.

Click here to verify bit.ly/[bank name]-App-VerifyFor (*Instruction*)

You can reschedule further delivery options by following here: [delivery service].parcel2ks.info (*Suggestion*)

MOVE 2

Problem statements

Problem statements occur in 72% of the texts. Their primary aim is to provoke action from the recipient. They can be presented as actual problems, potential problems, or issues or situations that might be seen as neutral.

We have noticed unusual activity on your internet banking (*Actual*)

Your subscription may have expired (*Potential*)

Your parcel is out for delivery (*Neutral*)

MOVE 3

Establishing credibility

The majority (65%) of the texts involve an attempt at establishing credibility of the sender as an authentic organisation or individual. This can be done by use of personal introductions and company names, providing details like account numbers and order references, or generic phrases commonly used in legal and standard marketing communications.

My name is Peter Walters

[BANK NAME] SECURITY ALERT

MEMBER ID #22102150

Terms and conditions apply

You may unsubscribe at any time

MOVE 4

Offers

Offers occur in 32% of the texts and are another strategy used to provoke action. They typically involve informing the recipient they have won something (money, goods) that they can now claim (usually by following a link).

$1000 [company] Gift Card is reserved just for you

Win a Tupperware Set worth £100

MOVE 5

Urgency statements

Urgency statements occur in 25% of the texts. They ask the recipient to act quickly on the Problem statement or Offer, either by explaining consequences for inaction, using urgent-sounding instructions, or expiry statements. Time frames tend to be abstract and avoid specific dates.

Package will be returned to sender if unpaid

GET STARTED NOW!

You only have 4 minutes 14 seconds to claim your prize!

While Solution and Problem statement are the most common moves in the current dataset, establishing credibility and Urgency statements are the only 'moves' reflected in the fraud literature, probably because these short-form texts represent a different type of fraud to others such as advance-fee scams, romance fraud and financial reports which are lengthier. The short texts analysed seem to follow the pattern of quickly introducing a statement designed to provoke action (e.g., Problem statement or Offer), followed by a solution (i.e. Your subscription has expired...click here to verify your account).

Other ways fraudsters use language in their communications

The analysis has so far revealed some of the language and communicative strategies that might be used by fraudsters. The researchers also uncovered three further features that could be an indication that a message is fraudulent.

Ambiguous messages

Fraudsters may try to draw recipients' attention to more than one piece of information, URL, or hyperlink.

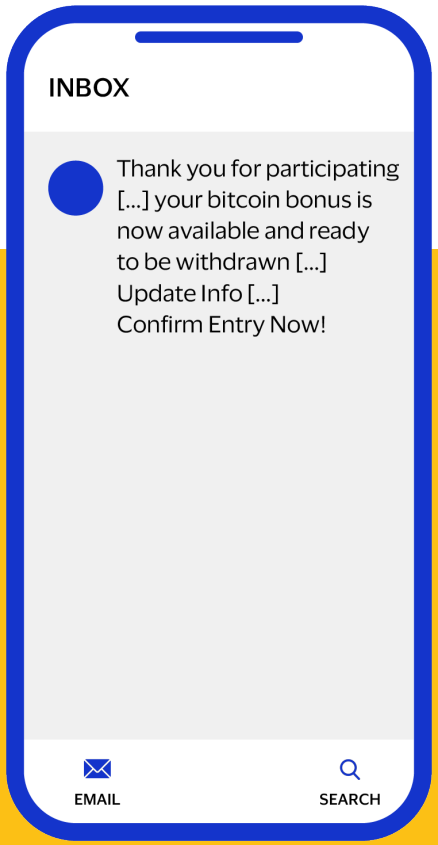
A mismatch between supposed sender and links

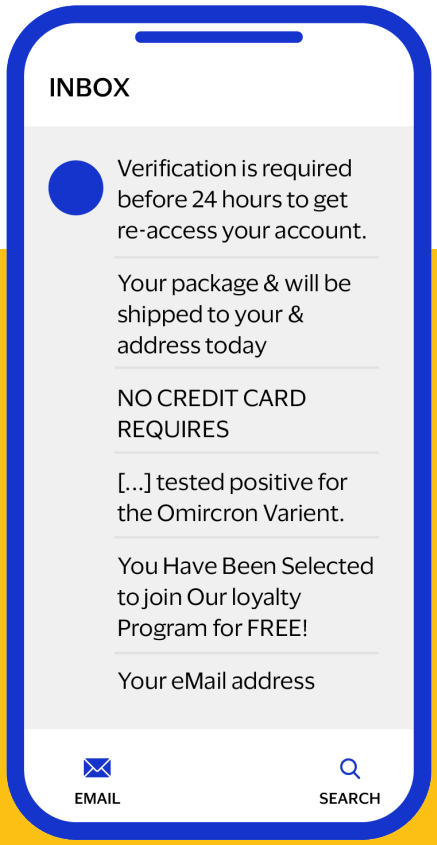
Fraudsters will make efforts to ensure URLs and hyperlinks look relevant to the source, and replicate colour schemes and styles of known brands and companies.

In these cases, it may be difficult to spot the signs of fraud. Although this isn't always the case, the researchers have observed a small number of cases where there is a clear mismatch between the supposed source and the URL.

Poor grammar, spelling, and typography

The researchers also found instances of errors in grammar and syntax – the arrangement of words and phrases – as well as misspelled words and random capitalisations.





Summary

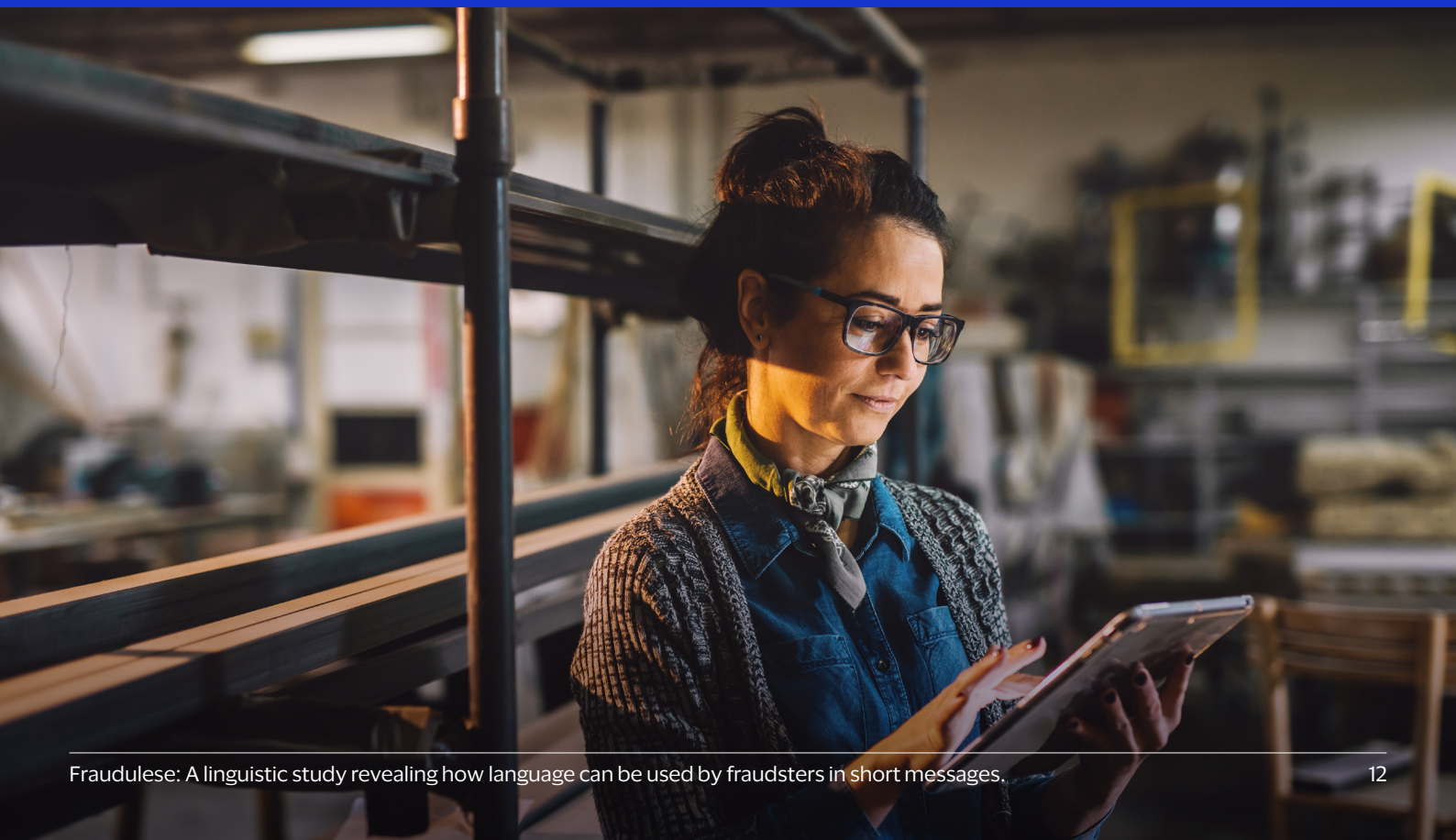
For the first time this linguistic study has revealed how language can be used by fraudsters in short, one-off messages.

Recognising the common words, phrases, and communicative strategies to look out for when receiving texts and emails that don't feel right will help consumers to spot the signs more easily and stay safe from fraud when shopping online.

With fraudsters increasingly using language techniques to make themselves appear credible, Visa is encouraging consumers to learn about 'Fraudulese' - the language of fraud, to help them pay with confidence.

The main findings of this exploratory study:

- **'Account', 'here'** and **'information'** are the most common words used by fraudsters in the dataset
- **'Click here', 'account information'** and **'gift card'** are the most common phrases identified in this research
- Often, fraudsters attempt to get consumers **to perform a particular action** and may prompt this by either suggesting a problem (e.g., a package is out for delivery) or making a tempting offer (e.g. suggesting a prize win)
- Fraudsters try to **convince the recipient of their credibility** in messages, sometimes using words and phrases that might be found in genuine communications.



References

[1] Research commissioned by Visa and conducted by Opinium with 2,001 nationally representative UK adults between 1st – 4th February 2022; Research commissioned by Visa and conducted by Opinium with 2,000 nationally representative UK adults between 18 March 2022 and 23 March 2022.

[2] Gov.uk. (2021) Public urged to protect themselves from online sales scams. Available at: https://www.gov.uk/government/news/public-urged-to-protect-themselves-from-online-sales-scams.\

[3] Hancock, J.T. et al. (2008) 'On lying and being lied to: A linguistic analysis of deception in computer-mediated communication', Discourse Processes, 45(1), pp. 1–23.

[4] This allows for drawing some tentative conclusions applicable to this exploratory dataset. Comparing this dataset to genuine communications would be extremely valuable in helping us to spot the features and strategies that are particularly characteristic of fraudulent texts.

[5] UK Finance (2021) Criminals exploit Covid-19 pandemic with rise in scams targeting victims online. Available at: https://www.ukfinance.org.uk/press/press-releases/criminals-exploit-covid-19-pandemic-rise-scams-targeting-victims-online.

[6] Bond, C.F. and DePaulo, B.M. (2006) 'Accuracy of Deception Judgments', Personality and Social Psychology Review, 10(3), pp. 214–234; George, J.F., Tilley, P. and Giordano, G. (2014) 'Sender credibility and deception detection', Computers in Human Behavior, 35, pp. 1–11; Blommaert, J. and Omoniyi, T. (2006) 'Email Fraud: Language, Technology, and the Indexicals of Globalisation', Social Semiotics, 16(4), pp. 573–605. doi:10.1080/10350330601019942; Lea, S., Fischer, P. and Evans, K.M. (2009) The psychology of scams: Provoking and committing errors of judgement, report for the Office of Fair Trading. Available at: https://webarchive.nationalarchives.gov.uk/20140402205717/http://oft.gov.uk/ shared_oft/reports/consumer_protection/oft1070.pdf; Freiermuth, M.R. (2011) 'Text, lies and electronic bait: An analysis of email fraud and the decisions of the unsuspecting', Discourse & Communication, 5(2), pp. 123–145; Carter, E. (2021) 'Distort, Extort, Deceive and Exploit: Exploring the Inner Workings of a Romance Fraud', The British Journal of Criminology, 61(2), pp. 283–302.

[7] Visa and Opinium research, March 2022.

[8] Visa and Opinium research, March 2022

[9] Gov.uk. (2021) Public urged to protect themselves from online sales scams. Available at: https://www.gov.uk/government/news/public-urged-to-protect-themselves-from-online-sales-scams.

Baker, P. (2009) 'The BE06 Corpus of British English and recent language change', International Journal of Corpus Linguistics, 14(3), pp. 312-337.

Blommaert, J. and Omoniyi, T. (2006) 'Email Fraud: Language, Technology, and the Indexicals of Globalisation', Social Semiotics, 16(4), pp. 573–605. doi:10.1080/10350330601019942.

Bond, C.F. and DePaulo, B.M. (2006) 'Accuracy of Deception Judgments', Personality and Social Psychology Review, 10(3), pp. 214–234.

Carter, E. (2021) 'Distort, Extort, Deceive and Exploit: Exploring the Inner Workings of a Romance Fraud', The British Journal of Criminology, 61(2), pp. 283–302.

Freiermuth, M.R. (2011) 'Text, lies and electronic bait: An analysis of email fraud and the decisions of the unsuspecting', Discourse & Communication, 5(2), pp. 123–145.

George, J.F., Tilley, P. and Giordano, G. (2014) 'Sender credibility and deception detection', Computers in Human Behavior, 35, pp. 1–11.

Gov.uk. (2021). Public urged to protect themselves from online sales scams. Available at: https://www.gov.uk/government/news/public-urged-to-protect-themselves-from-online-sales-scams.

Hancock, J.T. et al. (2008) 'On lying and being lied to: A linguistic analysis of deception in computer-mediated communication', Discourse Processes, 45(1), pp. 1–23.

Lea, S., Fischer, P. and Evans, K.M. (2009) The psychology of scams: Provoking and committing errors of judgement, report for the Office of Fair Trading. Available at: https://webarchive.nationalarchives.gov.uk/20140402205717/http://oft.gov.uk/ shared_oft/reports/consumer_protection/oft1070.pdf.

National Crime Agency (2022) Fraud. Available at: https://www.nationalcrimeagency.gov.uk/what-we-do/crime-threats/fraud-and-economic-crime.

Swales, J. (1990) Genre Analysis: English in academic and research settings. Cambridge: Cambridge University Press.

UK Finance (2021) Criminals exploit Covid-19 pandemic with rise in scams targeting victims online. Available at: https://www.ukfinance.org.uk/press/press-releases/criminals-exploit-covid-19-pandemic-rise-scams-targeting-victims-online.

DISCLAIMER

The data provided for this research consist of 155 short fraudulent texts (5,891 words), identified as 'phishing' and 'smishing' scams in the form of email and text messages. This allows for drawing some tentative conclusions applicable to this exploratory dataset. Comparing this dataset to genuine communications would be extremely valuable in helping us to spot the features and strategies that are particularly characteristic of fraudulent texts.