A WORK PROJECT PRESENTED AS PART OF THE REQUIREMENTS FOR THE AWARD OF A MASTERS DEGREE IN MANAGEMENT FROM NOVA SCHOOL OF BUSINESS AND ECONOMICS

INTEGRATION OF SPEECH-PROCESSING TECHNOLOGIES INTO ACTIVOBANK'S CLIENT INTERACTION PROCESS

Pedro Sebastião Ribeiro Estorninho da Mata Ribeiro, 1827

A project carried out with the supervision of Professor Paulo Faroleiro

January 2015

Abstract

This dissertation analyzes the possibilities of utilizing speech-processing technologies to transform the user experience of ActivoBank's customers while using remote banking solutions. The technologies are examined through different criteria to determine if they support the bank's goals and strategy and whether they should be incorporated in the bank's offering. These criteria include the alignment with ActivoBank's values, the suitability of the technology providers, the benefits these technologies entail, potential risks, appeal to the customers and impact on customer satisfaction. The analysis suggests that ActivoBank might not be in a position to adopt these technologies at this point in time.

Keywords: Speech-Processing Technologies, Retail Banking, Digital Banking,

Table of Contents

Abstract ..... **i**
Table of Contents ..... **ii**
1. Introduction **1**
1.1. Problem Statement and Approach 1
1.2. ActivoBank 1
1.2.1. The Multichannel Experience 2
2. The Technology **3**
2.1. Speech Recognition and Natural Language Understanding 4
2.1.1. Challenges and Limitations 4
2.1.2. Speech Recognition and Natural Language Understanding Applications 5
2.2. Speaker Recognition or Voice Biometrics 6
2.2.1. Challenges and limitations 7
2.2.2. Current Applications 8
3. Definition of potential uses for the technologies in ActivoBank **8**
3.1. Available Solutions 9
3.1.1. Voice Biometrics 9
3.1.2. Voice-enabled virtual Assistant 10
4. Speech-processing technologies in banking **11**
4.1. ING Netherlands - Inge 11
4.2. La Banque Postale – Talk to Pay 12
5. User Perception **13**
5.1. Perceived Benefits 14
5.2. Perceived Risks 15
5.3. Intention to Use 16
5.4. Methodology 16
5.5. Results 16
5.6. Discussion 17
6. Benefits **18**
6.1. Shareholders Perspective 19
6.2. Customers Perspective 20
6.3. Internal Perspective 20
6.4. Learning and Innovation 21
7. Risk Assessment and Controls **21**

7.1. Risk Assessment 22
7.2. Risk Treatment 23
7.3. Discussion 23
8. Analysis of Potential Suppliers 24
8.1. Evaluation of Potential Suppliers and Overall Analysis 24
9. Conclusions 25

1. Introduction

1.1. Problem Statement and Approach

A pioneer in Portuguese Mobile Banking, ActivoBank (AB) is always looking to strengthen its identity as forward-looking bank that makes use of technology to increasingly simplify its customers' lives. Making sense of the opportunity speech-processing technologies may represent, the bank is interested in understanding the possibilities of implementation. This dissertation is concerned with two main questions related with the introduction of these technologies in AB's client interaction process. First, considering the existent speech-processing technologies, which ones best suit AB? Second, should the bank consider incorporating these technologies in its offering?

To answer the first question an overall analysis of the speech-processing technologies will be conducted. Based on this, and considering the bank's strategic objectives the speech-processing technologies that best suit AB will be outlined. The following steps aim to answer the second question. First, an analysis of the different approaches to these technologies that have been used by FIs across the world will be conducted. Second, dimensions of customer experience that users perceive when judging the value of these technologies will be examined. Third, the nature of benefits associated with these technologies will be determined. Fourth, potential risks will be identified and options to manage them derived. Lastly, potential partners will be evaluated using a decision framework, to give AB the instruments in case it faces a supplier-selection decision.

1.2. ActivoBank

Founded in 1994, ActivoBank (then Banco7) was the first bank in Portugal dedicated to telephone banking. In 2010 the bank launched a rebranding initiative

introducing new offerings targeting people that embrace the use of new technologies and want both simplicity and convenience. The recent identity of the bank is grounded in five main principles: transparency, simplicity, trustworthiness, innovation and accessibility. The goal is to make customers' lives easier by delivering a simple but complete experience built upon a new concept of banking – Bank 2.0+[1] based in three main trends that are shaping today's world: the development of the mobile technology, the rise of social media and changing consumer behavior This is why AB maintains innovation at the core of most of its processes and products, expressed in the way costumers access the bank but also in the bank's communication strategy. For instance, the bank introduced a paperless opening account experience, replacing the paper signature with an iPad signature. The innovation culture of the bank has been widely recognized by national and international reviews.

1.2.1. The Multichannel Experience

Since the main objective of this work is to analyze how to integrate voice recognition technologies to improve the user experience of AB's customers while using remote banking solutions, it is important to describe in more detail the bank's website, mobile, and telephone channels.

Web Channel: The web channel offers AB's customers the possibility of managing their accounts 24 hours a day, seven days a week. In order to login, the client is required to provide its user identity code (obtained in an ATM), followed by three random numbers of a confidential seven-digit multichannel code. The navigation menu structured in a way that narrows the navigation options, enabling users to more easily and intuitively find content on the website.

[1] Concept inspired by the term Bank 2.0 - created by Brett King - that reflects the banks' evolution to a customer-centric strategy with a strong focus on technology and on the re-architecture of the core banking infrastructure.

Mobile Channel: In 2014 25% of AB's customers used a smartphone or tablet to conduct a transaction within the previous three months. To activate his account the user is required to (1) chose a four-digit PIN number, (2) provide its user identity code (this code can be used in one device only, meaning no one else can use that particular device to access mobile banking), (3) insert three random digits of the multichannel code, (4) enter a code he has received by SMS on his registered cell phone, and (5) enter again three random digits of the multichannel code. Once activated, each time the user logs in he is just required to enter the PIN code. AB offers two mobile banking apps available for Android and iOS devices: 'AB' and 'ActivoInvest'. The 'AB' app allows the customer not only to consult their assets and carry out bank operations, but also to constitute, reinforce their savings, or search for the branches' location. The 'ActivoInvest' app provides information on stock exchange indexes includes the most relevant business news and allows the customer to manage her investments portfolio.

Telephone Channel: The bank's call center handles outbound and inbound calls (40 agents in total). When calling, the client is required to dictate to the agent three digits of her multichannel code to ensure security is preserved. The security process can also require the clients to answer security questions in specific cases.

2. The Technology

The recent developments on the field of mobile devices, such as web-enabled smartphones, have resulted in new user expectations and needs in terms of services provided and ubiquitous availability (Schalkwyk, et al. 2010). This is especially true for the banking sector as users expect to be able to do banking irrespectively of the time, the place, or usage scenario (e.g. walking down the street). The operation of human-mobile interfaces such as touch screens or keyboards turns out to be inconvenient in

many occasions, reason why the development of speech-processing technologies may help to accommodate the rising user needs for ubiquitous mobile access.

Speech-processing technologies include Speech Recognition (SR) and Natural Language Understanding (NLU), and Speaker Recognition. The fundamental distinction between SR and NLU, and Speaker Recognition is that SR and NLU identify spoken words, i.e. they recognize what the person is saying, whereas the latter use features of a person's voice to ascertain the speaker's identity (Markowitz 2000).

2.1. Speech Recognition and Natural Language Understanding

Even though the complex technologies supporting SR systems vary much, the underlying technology of SR is essentially the same for all the major applications today. In the simplest sense, speech is input into a microphone or telephone, being the task of ASR systems to take this input and produce output as a string of words (Vimala and Radha 2012). Basically, when given with acoustic signal X, ASR systems decode it into the likeliest corresponding word sequence, W. Once the utterance is recognized, the next step is the understanding of the intent of the words (i.e. the semantic representation) using NLU. Finally, the system responds the user though voice output, text, or in the form of the requested action being performed (Rabiner and Juang 1993).

2.1.1. Challenges and Limitations

SR systems that support large vocabulary are increasingly being used as they become increasingly accurate. But even though progress has been notable, some challenges must be solved before widespread adoption can be fully realized. Most challenges relate to the technology's level of maturity and its ability to deliver a sufficiently robust, user-friendly and low-error performance. The difficulties arising from the casual nature of conversational speech and from acoustic environment distortion are two of the main technical challenges for SR technology (Deng and Huang

2004). The other - and probably the most relevant challenge - concerns usability. Conversational speech makes it more difficult to understand the spoken queries and return the correct output. This happens because vocabulary increases exponentially and input becomes much more unpredictable, increasing the complexity of the queries. The fact that noise conditions may vary enormously depending on the usage scenario also poses some problems to SR systems, which tend to perform better under controlled environments (Deng and Huang 2004). Still, over recent years SR technologies have become a lot more robust, improving the ability to filter out the surrounding noises and to successfully answer the queries. Since SR is still a challenging problem computationally, all the major consumer platforms that do it need to send the voice input to remote servers to recognize it and interpret it. This can cause SR on many devices to be slow, thus reducing the usability of the systems.

2.1.2. Speech Recognition and Natural Language Understanding Applications

The most familiar applications of SR and NLU systems today are in customer call center operations. NLU call routing solutions allow callers to speak to Interactive Voice Response systems (IVRs) in their one words without having to phrase their requests in pre-defined ways the system is expecting (Roukos 2008). The solution tries to understand the intent of the caller and connect him to the service he needs. But NLU and SR features can also be integrated with mobile apps through voice-enabled virtual assistants, enabling customers to perform routine operations using voice (e.g. Siri on iOS/iPhone, or Google Now on Android). Rather than tapping, or typing, the customers can get what they want by using their voice, as these virtual assistants understand the customers' conversation, reply to them, and carry out many tasks via voice commands.

2.2. Speaker Recognition or Voice Biometrics

When the task involves identifying the person talking rather than what is being said, the voice input must be processed to extract measures of speaker variability - such as physical characteristics associated with the vocal tract, or behavioral characteristics derived from pronunciation, gender, accents, etc. - as opposed to being analyzed by segments corresponding to phonemes or pieces of text (Peacocke e Graf 1990). There are two forms of speaker recognition (or voice biometrics): speaker verification and speaker identification. Since most forms of security based transactions (such as banking transactions) require a person to be verified rather than identified (Togneri and Pullella 2011), speaker verification will be explored in more detail compared to speaker identification. Verification systems have two steps: enrollment and verification. During the enrollment procedure the speaker's voice is recorded and stored in a database of reference templates. The verification phase begins with a claim of identity from the individual, such as entering a code or an accounting number, or allowing the system to access the ID of an input device linked to the user (Markowitz 2000) (such as a mobile app). The system then uses this information to retrieve the reference voiceprint and requests a voice sample from the claimant that is compared to the reference one. The results of the comparison process are then quantified and linked to an acceptance/rejection threshold to determine whether the two voiceprints are similar enough for the system to accept the claim (Markowitz 2000).

There are several different types of speaker identification: (1) text-dependent, (3) text-prompted, and (3) text-independent. Text-dependent systems request a prearranged password. Test-prompted systems ask the person to repeat a series of randomly selected digits or words. Finally, text-independent accept any spoken input (Markowitz 2000).

2.2.1. Challenges and limitations

Since the comparison results between the voiceprint produced during the verification and the reference voiceprint rely on an acceptance/rejection threshold, it becomes clear that verification systems can make two types of classification errors. They can either reject a speaker who is making an honest identity claim (false rejection), or mistakenly accept a speaker making a dishonest one (false acceptance). A challenge for banks using these systems is to define the actual point at which the trade-off between false rejection and false acceptance should be set. If the threshold is set too high, there will be few or no false accepts but a higher number of false rejects. If the threshold is set too low, the probability of mistakenly accepting a user who isn't whom he claims to be becomes higher. Finding the right balance between the false acceptance and false reject rates is a difficult decision and must take into account several factors such as the costs associated with false acceptance (reputational or operational) and customers' tolerance of false rejection, just to mention a few. One major area of concern for banks wanting to introduce speaker recognition systems is security, namely the susceptibility to spoofing attacks through the use of recorded voice. And even though it is fairly difficult to defeat a good speaker verification system with a recording (Markowitz 2000), some controls can be applied to ensure the system is not vulnerable to attacks. What's more, there is the risk that biometric data and biometric templates can be compromised during the enrollment phase when enrollment is poorly administrated (Dimitriadis and Polemi 2004), so it is vital to ensure a good enrollment process. Lastly, a challenge to attain efficient verification systems is the need to get clean audio. If by this point in the technology development issues such as authenticating users who have colds or as the ability to differentiate twins can be controlled, poor cell phone reception or extremely noisy environments can still be challenging (Felipe-Barkin 2011).

2.2.2. Current Applications

Voice biometrics can be deployed as an alternative to passwords, PINs and security questions to secure several self-service channels, including the web, mobile apps, and contact centers. For the web and mobile channels, voice biometrics can be used as a means of authentication and transaction validation. In the contact centers it can be used both as IVR authentication or authentication while contacting an agent.

3. Definition of potential uses for the technologies in ActivoBank

In order to define the ways in which speech-processing technologies can be integrated in AB, the first step is to outline the mission, or purpose behind it, and how it relates to the bank's strategic objectives.

As was previously mentioned, innovation is crucial to AB and is seen as an important driver of customer attraction and retention. This is reflected on a consistent investment in the digital channels, particularly the mobile banking platform, and on testing innovative ways of engaging customers. Traditional channels, such as physical branches and telephone are becoming less relevant as the digital channels rise. The number of customers using the mobile channel has been growing consistently since its introduction. In the first quarter of 2014 the number of mobile customers reached over 13,000 (~ 25% of the total number of customers) compared to only 9,000 the previous year – a 44% increase. This trend appears likely to continue as customer growth has been constant. But not only the number of customers using digital channels is growing tremendously, these channels are also less expensive, thus more attractive to the bank, than branch and telephone channels. A study conducted by Booz & Company argues that the cost per contact online represents less than 4% of the cost of a call handled by an agent (Booz & Company 2008). This is why AB is consistently trying to move more

transactions to self-service options such as mobile banking or the website. For these reasons, it becomes clear that a potential integration of speech-processing technologies in AB only makes sense in the context of the mobile and web channels.

3.1. Available Solutions

There are two uses for speech-processing technologies in the web and mobile channels: (i) voice biometrics for login authentication and transaction validation, and (ii) voice-enabled virtual assistant for the mobile app.

3.1.1. Voice Biometrics

As was mentioned before, voice biometrics can be deployed as an alternative to passwords and PINs to secure the web and mobile channels, and can be used as a means of authentication and transaction validation. For the web channel in particular, there are three ways the client can validate his identity with voice biometrics, and the bank should choose the alternative that fits the best with the bank's strategy and values. The more obvious solution is that the customer authenticates himself using the microphone from the computer. However, since clients visiting the web banking portal may be doing so from a computer without a microphone, the voice passphrase can be collected either by initiating an outbound call to the user's phone and asking the client to speak the passphrase, or by allowing him to access the mobile app (integrated with the website) and speak the password. By adopting this last approach, the bank is providing additional benefit of delivering a second authentication factor, further increasing security; however, using directly the computer's microphone may provide a faster experience.

Voice biometrics can be integrated with the methods of authentication already in place in a way that when trying to authenticate, if the user is below a pre-defined threshold, he is sent back to the traditional authentication mechanisms (e.g. he is required to type a PIN). The bank can also decide to use voice biometrics as an

additional layer within a multi-factor security architecture to strengthen security, independent of the circumstance, but considering AB current authentication procedures this would go against the bank's quest for simplicity. In a customer experience perspective, it would be better for AB to offer both login authentication and transaction validation capabilities as it would increase value of the service. Also, the bank should also give the customer the ability to decide whether he wants to use voice biometrics or traditional authentication methods in different circumstances (e.g. if the customer is driving he may opt to use voice biometrics, if he is in a crowded mall he might prefer a password). Another aspect worth considering is whether the passphrase is text dependent, text prompted, or text independent. To bring a unique experience to each customer, the best solution would be a text independent passphrase. The problem is that it requires longer samples for a voice print and is hard to use for login. To ensure security and still offer a good customer experience, text-dependent systems (e.g. the user says 'my voice is my password') or text-prompted appear to be reasonable options.

When introducing this technology, a bank must choose to host the solution on premises (on-site) or in the cloud. In hosted solutions, the voice biometrics system resides in the cloud, and the responsibility for the maintenance and management of the servers falls on the vendor. On-site voice biometrics, on the other hand, integrates with the bank's existing systems architecture, offering a higher level of security. Because data security is essential to any bank, on-site voice biometrics are a preferential choice.

3.1.2. Voice-enabled virtual Assistant

If AB chooses to introduce a voice-enabled virtual assistant to its customer service mobile apps, the way it can be done is by getting access to a platform that allows it to implement these functionalities directly within the apps. It is provided as a cloud-based solution as it makes use of remote serves for interpretation of the queries. Basically, as

the user performs voice commands, the voice input is sent to remote services for recognition and interpretation. With this technology, customers can speak through the smartphone microphone to the assistant. In response, the assistant talks or writes a message. For example, a user could ask "What was my current account balance yesterday?" The assistant would interpret that 'balance' refers to the euro amount in the current account, while 'yesterday' means to exclude transactions from today.

4. Speech-processing technologies in banking

Even though it is still not widespread, an increasing number of FIs is introducing speech-processing technologies into their self-service channels, namely the web and mobile channels. In order to support the study, banks that are already employing these technologies were examined. The objective is to gather the different type of approaches to speech-processing technologies that have been used by FIs, the methods they used to implement it, and the results they obtained. Several banks operating in different countries were identified (Table 1). Two of them (ING Netherlands and La Banque Postale) were analyzed more extensively as they provide two possibly interesting approaches to voice technologies. In Portugal no bank has deployed these systems.

Criteria	ING	Garanti	USBank	Barclays	ANZ	Santander	Unicredit	LBP
Assistant (mobile)	✓	✓	✓				✓	
Speech mining		✓						
Biometrics (mob)	✓*		✓		✓*			
Biometrics (website)								✓*
Biometrics (IVR)				✓				
Biometrics (agent)						✓		

***Not available yet**

Table 1 – Benchmarking summary

4.1. ING Netherlands - Inge

ING is a Dutch global FI offering banking, investment, and insurance services. In September 2014, ING Netherlands was the first bank in Europe to introduce voice

control features into its Mobile Banking App (ING First Bank in Europe to Provide Hands-Free Banking 2014). Inge is a voice-enabled virtual assistant that uses the capabilities of Nuance's Nina, a Software Development Kit (SDK) that enables NLU and TTS interfaces for mobile devices. With Inge, the two million users of the Mobile Banking App are able to speak via a human-like conversational interface to control the app. According to ING by adding a human aspect to their app, they are making mobile banking more personal (ING First Bank in Europe to Provide Hands-Free Banking 2014). Customers can retrieve their balance, make a transfer, or give a payment order using their voice. The app makes it possible to read out an account number instead of typing it, or to transfer money to a beneficiary in their address book. The person can then leave a voice description in the transfer note. But despite the many functionalities, Inge still presents some limitations when it comes to queries of a higher level of difficulty that require real-time availability of data and a higher degree of intelligence to derive the answers from specific dates. Some also argue that it still sounds a bit unnatural (Spelier 2014). ING is introducing hands-free banking in two steps. The first step, starting in mid-September let ING customers to start using Inge but requires them to use their PIN code to log on onto the mobile app and to authorize payments. The second step (scheduled to the end of 2014) will comprise authorization and logging on through voice recognition. Prior to the launch of the second phase, ING is inviting a group of customers to test voice recognition prior to its kick-off with the objective of gathering customer feedback that will be used to further improve this new mode before it is introduced to all customers.

4.2. La Banque Postale – Talk to Pay

La Banque Postale (LBP) - created in 2006 - is a subsidiary of La Poste, the French postal service. In response to the growing online payments fraud that threatens e-

business and banks, LBP launched the Talk To Pay project. The solution was created and designed by PW Consultants, a firm specialized in cards and payments. LBP worked closely with the firm since the beginning and took care of the compliance, risk and legal aspects. Basically, Talk To Pay is a security solution for online payments which includes voice biometrics authentication and the generation of dynamic visual cryptogram to allow payments. In some way, it is a two-factor authentication system as it is based on the smartphone as a personally owned device (the call originates from the system towards the user's phone) as well as on the unique voice characteristics of the person. With this this solution, the person simply picks up the phone and repeats a phrase to authenticate. The project began with a pilot phase that started in July 2013, after the authorization of the French CNIL to process biometrics data, and resulted in a prototype. The prototype will test if the bank can overcome what it considers to be the project's main challenge: the capacity to authenticate users in a wide range of use cases with a compromise between security and ease of use. In April 2014 several hundreds of workers and clients started testing the service in order to understand if the customer experience offered is acceptable. It is expected that once this experimentation phase ends, LBP will be providing Talk To Pay to all of its customers.

The projects analyzed are still very recent, which makes it difficult to appreciate what were the results, for instance in terms of customer satisfaction, potential revenues or brand image. This analysis may be completed in the future when there will be more data and information available.

5. User Perception

For the introduction of speech-processing technologies in AB's processes to be successful, users need to adopt them. By understanding what drives consumers to adopt

a new technology, AB is in a better position to evaluate the technologies from the consumers' perspective and see whether they perceive value from it. Byun & Byun (2013) highlight that in some situations consumers ignore new technologies because they perceive little benefits from its use. On the contrary, and according to Rogers (as cited by Byun and Byun 2013), if consumers perceive relative advantages from the use of a new technology, this may significantly influence the diffusion of an innovation. Another aspect worth noticing is that consumers' concerns toward using an innovation may impede its adoption, requiring an examination of the consumers' risk perception.

In order to examine the dimensions of customer experience, risks, and benefits that users perceive when judging the value of voice biometrics and voice-enabled virtual assistants, as compared to the use of traditional authentication methods and tapping, a survey instrument was developed for data collection. The questionnaire consisted of 29 questions to cover multiple variables of consumers' perceived benefits and risks. The variables of perceived benefits identified consisted of perceived ease of use, usefulness, convenience, and increased security (for voice authentication). The identified dimensions of perceived risks were reliability risk, and data and privacy risk. These dimensions were extracted based on a literature review.

5.1. Perceived Benefits

Davis (1986) developed a Technology Acceptance Model (TAM) in which he identified **perceived ease of use** has an attribute of innovation that influence adoption (Moore and Bensabat 1991). Since the use of voice for biometrics and voice commands seems natural and appropriate, it is expected that consumers will perceive the considered technologies as 'easy to use':

H_1-1: Consumers perceive a voice-enabled virtual assistant to be easy to use.

H_1-2: Consumers perceive voice biometrics to be easy to use.

The importance of **perceived usefulness** (also introduced by Davis in 1986) has been extensively recognized in the field of e-banking (Jahangir and Begum 2008). According to Kulviwat, et al. (2007), it refers to the degree to which a person believes that using a technology will benefit her in performance of some task. Davis et al. (as cited by Kulviwat, et al. 2007) suggested that the perceived usefulness is an important factor in determining not only the initial acceptance of an innovation but also its adoption and continued use. In this respect, the following hypotheses were formulated:

H_2-1: Consumers perceive a voice-enabled virtual assistant to be useful.

H_2-2: Consumers perceive voice biometrics to be useful.

The benefits of **convenience** are enabling factors that make it easy for the consumer to do business with a bank, and have an impact on his level of satisfaction (Gupta and Dev 2012). Two hypothesis are framed:

H_3-1: Consumers perceive a voice-enabled virtual assistant will increase convenience.

H_3-2: Consumers perceive voice biometrics will increase convenience.

Security is essential to every transaction and especially in the banking sector where clients trust their money based on promises done (Koksal and Dema 2014). Some banks employing voice authentication systems argue these systems are secure and reduce the chances of fraud (Sambidge 2014). Based upon this, the following hypothesis has been developed:

H_4-1: Consumers perceive voice biometrics will increase security.

5.2. Perceived Risks

The use of SR systems in banking applications may raise questions concerning **data and privacy protection**. Consumers may be concerned about whether there are security measures taken to protect their data and what kind of information is stored. Thus, it is

expected that the security and privacy of their information would be one of the risks consumers perceive when using a voice-enabled virtual assistant:

H_5-1: Consumers are concerned about security and data privacy, when using a voice-enabled virtual assistant.

5.3. Intention to Use

Another aim of the survey was to understand to which measure customers are willing to adopt and use the innovations.

5.4. Methodology

Participants for this study were individuals who operate current or savings account in AB, age 18 and above, in Portugal. There were 108 valid answers (64 male respondents and 44 female). Individuals aged 24 to 40 years old represent 59% of the total. Individuals aged 41 to 60 years old account for 38% of the sample. Only 3% were between 18 and 24 years and 1% over 60 years old. A five-point Likert scale ranging from a scale of 1 being 'Strongly disagree' to 5 being 'Totally agree' was used.

5.5. Results

Hypotheses $_1$-1 and $_1$-2: The respondents are asked if they perceive the assistant and voice biometrics as easy to use. Over 74% of respondents say they feel a virtual assistant would be 'Easy' (50%) or 'Very Easy' (24,53%) to use and the mean is 3.83 (out of 5), suggesting the null hypothesis cannot be rejected. The phenomenon is not as evident when it comes to voice biometrics, but it is still relatively similar. 50.96% of the users consider it would be 'Easy' and 16.35% 'Very Easy', and the mean is 3.52.

Hypotheses $_2$-1 and $_2$-2: The answers are very similar for both the voice assistant and voice biometrics (mean of 3.73 in one case and 3.61 in the other). Approximately 50% of the respondents (52% vs.50% respectively) partially agree that the technologies

are useful, 17% consider the assistant as very useful, while 20% consider voice biometrics very useful. The percentage of individuals with no opinion is 10% and 16%, whereas the percentage who do not consider the technologies useful is 21% (7% totally disagree, and 14% partially disagree) and 14%, respectively (3% and 11%).

Hypotheses $_3$-1 and $_3$-2: Approximately 42% of the respondents (41% vs. 43% respectively) agree that would help saving time, 17% consider the assistant as very helpful in these terms, while 20% consider voice biometrics. The percentage of individuals with no opinion is 22% and 14%. The percentage of individuals who think otherwise is composed of 6% who totally disagree, and 14% partially disagree and, respectively, 9% and 18%.

Hypothesis $_4$-1: 13% of the clients consider that definitely voice biometrics will not increase security, another 27% do not perceive any increase in security; 29% have no opinion, and only 31% perceive any increase in security (25% partially agree, 6% totally agree).

Hypothesis $_5$-1: Close to 53% of the customers express concerns over security (10% totally agree that a voice assistant can endanger security and compromise their privacy, while 43% partially agree); 21% remain neutral.

5.6. Discussion

Customers perceive benefits overall, both for the assistant and voice biometrics, which may positively influence consumer's adoption. However, they reveal some concerns about security and data privacy in the virtual assistant, and are skeptical in what concerns the security of voice biometrics. This situation is expected to have a negative effect on adoption. This suggests that the adoption can only happen if the consumers are well-informed not only to the relative advantages of these technologies,

but about the security they provide. Further research on consumers may be positive as the sample may not reflect the population in the best possible manner.

6. Benefits

Projects should yield value to the organization. A framework to determine the nature of the benefits associated with projects is the balanced scorecard (BSC). The BSC provides a structure that can be of help when thinking through and identifying possible benefits that might result from a project. By using the BSC, organizations do not have to rely only on financial measures as the sole benefits' indicators. The BSC links four perspectives on business projects that overall combine to provide the value proposition from the project. These are:

Perspective	**Description**
Shareholders	What are the expected financial benefits during the project's lifetime?
Customers	What are the benefits they will receive with the project? Will the project enable the bank to work with a different market of customers?
Internally	How will the project contribute to the improvement of the bank's business processes?
Learning & Innovation:	How much is the organization expected to learn from the project? And how does the innovation compare to competitors?

Table 2 – BSC Perspectives

A series of expected benefits were listed and explained according to the different perspectives (Table 2). Benefits were separated between the ones coming solely from the adoption of voice biometrics, those exclusive to a voice assistant, and the benefits that are common to both solutions (e.g. customer experience is expected to improve the bank decides to launch voice biometrics or a voice assistant only).

	Benefits		
Perspective	*Voice Biometrics (VB)*	*Voice Assistant (VA)*	*VB + VA*
Shareholders			Cost savings due to a lower amount of calls to the contact center. New clients.
Customers	Increased security, which will positively impact customer satisfaction.		Improved customer experience. Increased awareness of the AB brand.
Internal			Reduction of calls to call center, allowing the bank to focus more on

	the outbound activity.
Learning & Innovation	Systems and technology learning. Reinforcement of the position as an innovative bank.

Table 3 – List of the potential benefits

6.1. Shareholders Perspective

The identification of benefits for the shareholders is particularly challenging as the benefits are not purely financial and is difficult to derive direct monetary benefit from speech-processing technologies on the web and mobile channels (the service is not directly charged to the customer); yet it would be foolish not to consider that there can be indirect financial benefits coming from these solutions. The introduction of these technologies may result in a slight increase in the number of customers, attracting more sophisticated customers that privilege self-service channels and value innovation. But while this benefit is not expected to be much relevant, the cost savings the bank may achieve due to a reduction on calls to the contact center may be relevant. A recent study by Research Now (2013) revealed that 57% of all calls into the contact center are generated by individuals that first attempted to use self-service channels. One of the key reasons for this is the failure to authenticate within self-service channels (they are unable to access because they forget their password, they can make mistakes while typing password codes, etc.). With voice biometrics the customer is only required to use her voice to authenticate, thus reducing the costs of servicing clients through contact center agents that would have otherwise used self-service. The same reasoning applies for the voice-enabled virtual assistant; clients tend to rely on one self-service channel for their day-to-day banking activities, and leverage other channels when the preferential channel is unable to address the customer's request. Sometimes customers are not able to find the information they are looking for using the website or the mobile app, calling the contact center with questions. By providing information on the

individual's banking activities and the bank's products in an intuitive manner, thus improving the self-service experience, the voice-enabled virtual assistant can help shift calls away from agents, reducing operating costs.

6.2. Customers Perspective

The use of voice biometrics can enhance security if coupled with other authentication methods, thus increasing costumers' sense of trust. Customer experience can also be improved using voice biometrics systems. Instead of PIN numbers that can be forgotten or skimmed, voice is an inherent characteristic of the person, meaning that the client does not need to memorize several codes (reduces cognitive load). Both technologies contribute in other ways to improve customer experience, especially because they increase convenience and ubiquity. The possibility of hands free usage, allows the customer to conduct is day-to-day banking virtually everywhere at any time, including in otherwise impractical usage scenarios (e.g.: while driving). Another benefit coming from these solutions is that they may contribute to increase AB's brand awareness. AB, being a small bank does not have very high brand awareness. It is natural that if AB introduces speech-processing technologies into its client-interaction processes, the bank will receive airplay on televisions, and attention on digital media or specialty magazines for some time benefiting from the fact of being the first one in Portugal to do so. This happened when AB introduced the paperless account opening experience, and so it is predictable that this can be repeated. A substantial increase in brand awareness is expected during and following this period.

6.3. Internal Perspective

As it was previously mentioned, it is expected that the introduction of speech-processing technologies may help prevent avoidable calls. This might, in turn, help the contact center to move away from a service mindset toward one focused in sales. By

reducing the number of calls, the number of inbound agents may be reduced and this can engage in outbound activities, increasing cross-selling and customer loyalty.

6.4. Learning and Innovation

If AB decides to introduce speech-processing technologies it will presumably be the first bank to do so in Portugal. This will reinforce AB's position as an innovative bank in the Portuguese context. There is also another benefit which is the learning over systems and the technology that will come if the bank adopts the technologies. The accumulated knowledge can also serve to facilitate the introduction of other speech-processing technologies in the future (e.g. call steering).

In conclusion, there are several benefits in several dimensions for both voice biometrics and the voice-enabled virtual assistant. But if it is true that most benefits may result from the adoption of either of the solutions, they are enhanced if both are adopted simultaneously, i.e. the benefits will likely be more significant in case both technologies are launched as a whole. It should be noted, however, that voice biometrics brings benefits that are exclusive to the solution, namely in terms of security, meaning that the solution can be viable by itself.

7. Risk Assessment and Controls

Projects always carry some degree of uncertainty. For instance, there may be uncertainty on how the customers will respond, about the costs, etc. The consideration of risk is one aspect on managing uncertainty (Maylor 2010). According to Maylor (2010) an evaluation or risk is important as it may show whether or not a venture is worth pursuing. The objective of this section is to identify and derive options to manage the risk associated with voice biometrics and the virtual assistant technologies.

The procedure adopted in this work is based on the approach proposed by Cooper et al (2005). It consists of five stages: (i) establishing the context (refer to Appendix 1 to consult the set of the criteria of the project), (ii) identifying the risks (refer to Appendixes 2 and 3 for an identification of the risks), (iii) analyzing the risks, (iv) evaluating the risks, and (v) treating the risks (Cooper, et al. 2005). It should be noted that a quantitative assessment was concluded to be difficult and based on data that could not be suitably supported, and so a simpler qualitative approach (based on descriptive scales) was adopted. Plausible risk scenarios were developed. The consequences and likelihood of the potential risks were explored to define qualitative estimates of the level of risk. Finally, based on these estimates, treatment actions were designed.

7.1. Risk Assessment

Tables 8 and 9 present the level of risk for each of the technologies, together with the predictable consequences of the risks and the likelihood of occurrence, giving a comprehensive assessment of the risks of these solutions.

Risks	Consequences	Likelihood	Level of risk
VA1: The cloud provider fails.	Minor (customers unable to use the service temporarily)	Rare	Low
VA2: The technology is not accurate	Major (affects adoption of the technology and bank's image)	Possible	High
VA3: Security risks associated with the cloud.	Major (affects the bank's reputation and trust)	Rare	Medium
VA4: More time than predicted to launch.	Moderate (timing can be compromised, affecting success)	Likely	Medium
VA5: Customers do not adopt the technology.	Major (severely affects the success of the deployment)	Possible	High

Table 4 –Voice-enabled virtual assistant risk assessment

Risks	Consequences	Likelihood	Level of risk
VB1: The technology is not accurate	Major (affects adoption of the technology and bank's image)	Possible	High
VB2: Spoofing attacks	Major (affects the bank's reputation; can bear high costs)	Unlikely	Medium
VB3: Security of the data bases.	Major (affects the bank's reputation; can bear high costs)	Rare	Medium
VB4: More time than predicted to launch	Moderate (timing can be compromised, affecting success)	Likely	Medium

VB5: Customers do not adopt the technology.	Major (severely affects the success of the deployment)	Possible	High

Table 5 –Voice biometrics risk assessment

7.2. Risk Treatment

Possible risk treatment options were generated for the Medium and the High risks (refer to appendix 4. The risk treatment options were prompted in the areas of prevention and mitigation to aid in generating a comprehensive list. Options are examined in terms of their effectiveness and cost. It should be noted that some risks will benefit from the risk treatment options recommended for other risks.

7.3. Discussion

The majority of the risks identified have either a 'Medium' or a 'High' level of risk. This suggests that in case the bank introduces these technologies at this point in time, it will carry sizeable risks. If it is true that some of the High risks (VA2 and VB1) are easily preventable in case the bank performs independent testing (in case non of the vendors passes the test, that should be a big obstacle to the implementation). The risk that customers do not adopt the technology can be difficultly preventable and influenced, meaning AB should be careful with this matter.

The outcomes from the previous qualitative analysis are sufficient to give some insights to AB on the necessary decisions about the risks of this kind of solutions. However, in case the bank decides to adopt the technologies, the development and implementation of a Risk Action Plan is recommended and a more detailed analysis of the advantages and disadvantages (as well as costs and benefits) of each risk treatment option may be necessary.

8. Analysis of Potential Suppliers

AB does not have the internal capabilities to deliver speech-processing technologies on its own. The technology must be provided by external firms. Therefore, it is important to commit considerable effort seeking and selecting suitable partners and to really understand what are the costs involving each option. To assist AB, a multi-criteria decision framework consisting of supplier characteristics (e.g. location) and bid/product characteristics (e.g. quality) was developed, potential partners were evaluated in light of these criteria, and an overall analysis was conducted.

8.1. Evaluation of Potential Suppliers and Overall Analysis

Approximately fifteen companies were contacted in order to assess if they had the required technologies. Of those, three companies (Company A, B and C)[2] had the technology and provided information on the cost of the solutions. Company A specialized only on Voice Biometrics (for both mobile and the website platforms, and with the possibility of extending to contact centers), Company B offered a voice-enabled virtual assistant, Company C had both technologies (for mobile). Overall, the three companies have different strengths, but also reveal some weaknesses. Company A has great reputation and has the advantage that its platform is common to all the channels. Company B's solution is expensive and some concerns remain on whether it will work properly in Portuguese. The fact that Company C provides both a voice biometrics solution and a voice assistant solution can result in theory in an easier integration, and also reduce internal development and overhead costs, however, the fact that the price is based on number of licenses and not the number of actual users is a

[2] It is not possible to disclose the companies' names so they will be referred to as Company A, Company B and Company C.

negative aspect[3]. The price (refer to Appendix 5) of the voice biometrics solution, for Company A can vary between €61,750 in a low adoption scenario and €69,250 in a high adoption scenario (already including an estimate of internal development costs). Assuming 30,000 users (ie. the number of downloads of the app), the voice assistant solution from Company B would cost approximately €180,000, including server and maintenance costs for a two-year period. Assuming the same 30,000 users, the total cost of Company C's solutions would be close to €212,000, to which should be added €26,280 annually (starting in year 2) for maintenance and upgrade (refer to Appendix 6). Even though the prices of the three companies are relatively affordable (if we consider other projects in banking), it remains unclear, however, if a bank with AB's size is in a position to undertake an investment of this magnitude. The bank is on the way to turn profit for the first time in three years, by the end of the year, and for that reason it cannot afford to invest money in solutions without being sure of its value. For a bank with bigger dimensions this would not stand as a problem. The fixed costs of the solutions (e.g. licensing) would be less relevant considering the overall amount of money spent by the bank and the variable costs would directly depend on the success of the technology (i.e. user adoption), meaning that in case the technologies turned into a failure they would not impose significant costs.

9. Conclusions

This dissertation was set out to explore the possibilities of integration of speech-processing technologies into ActivoBank's client interaction process. The study has offered an evaluation perspective on different topics relevant to answer the research questions. The preferential channels in which these technologies could be applied to AB

[3] Even though in theory the number of licenses equals the number of downloaded apps in total, ActivoBank can think of possibilities to reduce this risk (see *Appendix 4)*.

were concluded to be the website and the mobile channels, as they are in accordance with AB's strategy to reduce interaction through the telephone channel, and have witness an increase in usage rates over the last years. The analysis also suggests that for these channels, there are two uses that fit AB's vision: voice biometrics for login authentication and transaction validation, and voice-enabled virtual assistant for the mobile app. These two technologies were analyzed through different angles allowing some conclusions. (1) Even though FI are increasingly introducing these technologies in different countries, it is not a large-scale phenomenon. It is not possible to know if this tendency will continue. Also, the cases studied also do not permit to infer on the relative success of these initiatives as they are very recent; (2) Customers perceive relative advantages overall. However, they reveal some concerns about the security and data protection. These two aspects suggest that these adoption can only happen if the consumers are well-informed not only to the relative advantages of these technologies, but especially about security; (3) These technologies potentially bring benefits both to the customers and the bank itself, especially if it is the first to introduce them in the country. However, the benefits for the bank are dependent on diverse factors, requiring future analysis; (4) Introducing these technologies carry risks that at this point in time, can potentially be difficult to cope with as the technologies are still not in the maturity phase; (5) the costs are considerably high if we consider AB's size and its reality, but the low number of proposals is a limitation of this dissertation.

These conclusions suggest that the advantages AB would take from investing in these technologies now would not compensate the associated costs and risks from this decision. AB should nonetheless follow the developments of these technologies in terms of their penetration in the market and results in other countries, assuring they do not lag behind the competition if the tendency

References

Abdullah, Moha A., Noor H. A. Manaf, Muhammad-Bashir O. Yusuf, Kamrul Ahsan, and Ferdous S.M. Azam. "Determinants of Customer Satisfaction on Retail Banks in New Zealand: An Empirical Analysis Using Structural Equation Modeling." *Global Economy and Finance Journal* 7 (1) (2014): 63-82.

Booz & Company. "Redefining the Mission for Bank's Call Centers: Cut Costs, Grow Sales, or Both." 2008.

Byun, Sookeun, and Sang-Eun Byun. "Exploring Perceptions Toward Biometric Technology in Service Encounters: A Comparison of Current Users and Potential Adopters." *Behaviour & Information Technology* 32 (3) (2013): 217-230.

Cooper, Dale, Stephen Grey, Geoffrey Raymond, and Phil Walker. *Project Management Risk Guidelines: Managing Risk in Large Projects and Complex Procurements.* John Wiley & Sons, Ltd, 2005.

Deloitte Center for Banking Solutions. "Evolving Models of Retail Banking Distribution." 2008.

Deng, Li, and Xuedong Huang. "Challenges in Adopting Speech Recognition." *Communications of the ACM* 47 (1) (2004): 69-75.

Dimitriadis, Christos K., and Despina Polemi. "Biometrics - Risks and Controls." *Information Systems Control Journal* 4 (2004).

Felipe-Barkin, Eric. "Tips for a Successful Voice Biometrics Deployment." *Speech Technology Magazine,* November 1, 2011.

Gao, Yuqing, Bowen Zhou, Weizhong Zhu, e Wei Zhang. "Handheld Speech to Speech Translation System." In *Automatic Speech Recognition on Mobile Devices and over Communication Networks*, de Zheng-Hua Tan e Borge Lindbeg, 327-345. Springer Sciene+Business Media, 2008.

Gupta, Aayushi, and Santosh Dev. "Client Satisfaction in Indian Banks: An Empirical Study." *Management Research Review* 35 (7) (2012): 617-636.

"ING First Bank in Europe to Provide Hands-Free Banking." *Mobey Forum.* September 15, 2014. http://www.mobeyforum.org/ing-first-bank-in-europe-to-provide-hands-free-banking/ (accessed November 21, 2014).

Jahangir, Nadim, and Noorjahan Begum. "The Role of Perceived Usefulness, Perceived Ease of Use, Security and Privacy, and Customer Attitude to Engender Customer Adaptation in the Context of Electronic Banking." *African Journal of Business Management* 2 (1) (2008): 32-40.

Koksal, Yuksel, and Oneda Dema. "An Investigation of the Essential Factors on Customer Loyalty in Banking Sector: A Case of Albanian Retail Bank." *Journal of Management and Economics* 21 (1) (2014): 357-368.

Kulviwat, Songpol, Gordon C. Bruner II, Anand Kumar, Suzanne A. Nasco, e Terry Clark. "Toward a Unified Theory of Consumer Acceptance Technology." *Psychology & Marketing* 24 (12) (2007): 1059-1084.

Markowitz, Judith A. "Voice Biometrics." *Communications of the ACM* 43 (9) (2000): 66-73.

Maylor, Harvey. *Project Management.* Pearson Prentice Hall, 2010.

Moore, Gary C., and Izak Bensabat. "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation." *Information Systems Research* 2 (1991): 192-222.

Peacocke, Richard D., e Daryl H. Graf. "Introduction to Speech and Speaker Recognition." *Computer*, 1990: 26-33.

Rabiner, Lawrence, and Biing-Hwang Juang. *Fundamentals of Speech Recognition.* Prentice-Hall International, 1993.

Research Now. "Web Self-Service Preferences Survey." 2013.

Roukos, Salim. "Natural Language Understanding." In *Springer Handbook of Speech Processing*, by Jacob Benestry, M. Mohan Sondhi and Yiteng A. Huang, 617-626. Springer Berling Heidelberg, 2008.

Sambidge, Andy. "Abu Dhabi Bank Says to Boost Security with Voice Biometrics." *Arabian Business.* October 19, 2014. http://www.arabianbusiness.com/abu-dhabi-bank-says-boost-security-with-voice-biometrics-568433.html; (accessed November 10, 2014).

Schalkwyk, Johan, et al. ""Your Word is my Command": Google Search by Voice: A Case Study." In *Advances in Speech Recognition*, de Amy Neustein, 61-90. Springer US, 2010.

Şen, S., Başligil H., Şen C.G., and H. BaraÇli. "A Framework for Defining both Qualitative and Quantitative Supplier Selection Criteria Considering the Buyer-Supplier Integration Strategies." *International Journal of Production Research* 46 (7) (2008): 1825-1845.

Spelier, Pascal. "Inge: Eerste Stap in Bankieren Met Stembediening." *Marketing Facts.* September 11, 2014. http://www.marketingfacts.nl/berichten/inge-eerste-stap-in-bankieren-met-stembediening (accessed November 21, 2014).

The Boston Consulting Group. "Distribution 2020: The Next Big Journey for Retail Banks." 2013.

Togneri, Roberto, and Daniel Pullella. "An Overview of Speaker Identification: Accuracy and Robustness Issues." *IEEE Circuits and Systems Magazine*, 2011: 23-61.

Vimala, C, and V Radha. "A Review on Speech Recognition Challenges and Approaches." *World of Computer Science and Information Technology Journal* 2 (1) (2012).

APPENDIX

Criteria	Notes
Performance	Performance, availability and reliability as required.
Image and reputation	Customers support and trust. AB seen has an innovative bank that brings value to the customer.
Costs	Costs should be adapted to the reality of the bank.
Timing	Development and implementation within a reasonable timeframe.
Security	The highest standards of security must be maintained.

Appendix 1 – General requirements involving speech-processing technologies

Main issues	Risks	Description
Performance	The cloud provider fails.	The cloud service provider is unexpectedly brought down temporarily, thus rendering the use of the assistant impossible.
	The technology is not accurate	The technology fails to achieve the desired speech recognition accuracy and speed in a real-world context creating frustration among customers.
Security	Security risks associated with the cloud.	Attackers can compromise confidentiality and integrity of the data which cause harm to the bank's reputation.
Timing	Development phase takes more time than predicted.	Development phase requires the work of specialists and access to data. The time needed to complete this phase can be unpredictable.
	Delay in the project.	Difficulties in integrating the solution to the bank's systems delay the project.
Customer	Customers do not adopt the technology.	A small percentage of the customers use the service. This brings several problems: if the licensing fees are based on the number of downloaded apps, it means the bank is paying for a service only a handful of customers use; it will mean that the technology brings no value to the bank.

Appendix 2 – Risk identification and description (voice-enabled assistant)

Main issues	Risks	Description
Performance	The technology is not accurate.	The technology fails to achieve the desired accuracy (FTER) and speed in a real-world context creating frustration among customers and reducing trust in the solution.
Security	Spoofing attacks. Security of the databases.	The technology is susceptible to spoofing attacks. Computers and networks used in the authentication process not being secure.
Timing	(See Table 4)	(See Table 4)
Customer	(See Table 4)	(See Table 4)

Appendix 3 – Risk identification and description (voice biometrics)

Risks	Prompt	Possible option	Evaluation
VA2	Prevention	Perform independent testing executed within the bank's own environment.	(High Effectiveness, Medium Cost)
VA3	Prevention	Use auditors to examine the controls at the cloud service provider, their commitment to following up on potential security incidents and ability to respond promptly to occurrences.	(High Effectiveness, High Cost)
VA4	Prevention	(External) Share the risk with the vendor when contract is negotiated ensuring the vendor is able to develop the solution on time ('contract failure risk').	(Medium Effectiveness, Low Cost)
VA5	Prevention	Ensure the benefits are clearly stated and the customer is aware of them through marketing campaigns.	(High Effectiveness, High Cost)
VA5	Prevention	Try to negotiate a contract in which the licensing fee is based on the number of users and not the number of licenses.	(High Effectiveness, High Cost)
VB1	Prevention	Perform independent testing executed within the bank's own environment and user population guidelines to understand actual performance.	(High Effectiveness, Medium Cost)
VB2	Mitigation	Adjust the probability thresholds to increase security; consider additional layers of security.	(High Effectiveness, Low Cost)
VB3	Prevention	Ensure that the computers and networks used in the authentication process have been secured and are being monitored on an ongoing basis.	(High Effectiveness, Low Cost)
VB4	Prevention	(See VA4)	(See VA4)
VB5	Prevention	(See VA4)	(See VA4)

Appendix 4 –Risk treatment options for the Medium and High risks

	Low Adoption	Medium Adoption	High Adoption
Platform price	€18,000	€18,000	€18,000
Number of users	7,500 (~25% of total)	15,000 (~50% of total)	22,500 (~75% of total)
Unit Price* (decreases the more the users, but not for this scale)	€0,50	€0,50	€0,50
Total usage price	€3,750	€7,500	€11,250
Development costs		€40,000*	
Total Cost	**€61,750**	**€65,500**	**€69,250**

*the UP decreases the more users are in the platform, but only for higher values
*it is calculated as an average of maximum and minimum values suggested

Appendix 5 – Price of voice biometrics solution, for Company A (three scenarios)

Product	Number of Licenses	Unit Price	Total Price
Mobile Speech Recognition SDK	30,000	€1,46	€43,800
Mobile Voice Biometrics SDK	30,000	€1,46	€43,800
Mobile Text-to-speech SDK	30,000	€1,46	€43,800
Project Management		€40,572	
Development Costs		€40,000	
Total Costs		**€211,972**	
Annual Maintenance & Upgrade*		**€26,280**	

*The quote was given in dollars; it is converted to EUR at the currency rate as of December 18, 2014
*20% of the total license price (€131,400); 1st year's M&U is included in the price

Appendix 6 – Price of voice biometrics and virtual assistant solutions (Company C)

OTHER APPENDIX (COMPLEMENTARY)

User Perception: Complementary Information

Construct	Measurement Questions
Ease of use	It would be easy to use a voice-enabled virtual assistant.
Convenience	Compared to tapping/typing, the use of a voice-enabled virtual assistant would: … be more convenient. … … make my banking transactions less time consuming. … would allow me to do my day-to-day banking in circumstances otherwise impossible (e.g. while driving).
Performance / Reliability	Q 23would perform as it is supposed
Information Privacy Risk	The use of a voice-enabled virtual assistant makes me concerned about security and data privacy.
Usefulness	The use of a voice-enabled virtual assistant would prove to be useful.
Intention to Use	Q 18

Appendix AA– Constructs and measurement questions from the survey (Voice Assistant)

Construct	Measurement Questions
Ease of use	It would be easy to use voice authentication.
Convenience	Compared to authentication methods in place, the use of voice authentication would: … be more convenient. … make my banking transactions less time consuming. … would allow me to do my day-to-day banking in circumstances otherwise impossible.
Usefulness	The use of voice authentication technology would prove to be useful.
Increased Security	… improve the level of security.

Performance / Reliability	Q 29
Intention to Use	Q 26

Appendix AB – Constructs and measurement questions from the survey (Voice Biometrics)

Risk Assessment and Controls: Complementary Information

Establishing the Context

Establishing the context is an important step to define the key elements for structuring the risk identification and assessment process (Cooper, et al. 2005). During this phase, the requirements of the company and the key stakeholders are used to derive a set of criteria for the projects, which will then be used to determine scales against which the consequences of risks will be assessed (Cooper, et al. 2005). The following table shows the list of criteria, or general requirements, of a potential project involving speech-processing technologies.

Risk Identification

The risk identification process aims to predict the key risk outcomes (Maylor 2010), i.e. indicators that could affect the objectives. For this work, two sources were used to identify the potential risks. One was brainstorming with AB's employees that work on the areas of innovation and IT. The other was collection and analysis of opinions from the stakeholder community, such as customers and the technology providers.

Risk Assessment

The purpose of risk assessment is to develop priorities for the previously identified risks, which will be used to determine which of the risk events should be treated as it is unlikely that all risks can be managed (Cooper, et al. 2005). This process will determine how likely specific events are to occur and the magnitude of their consequences.

First, a five-point descriptive consequence scale was developed (Appendix B). This scale gives a view of the overall consequences of the risks to the established criteria.

Rating	Potential impact on criteria
A Critical	Most criteria threatened, several severely affected, compromising the success of the project and damaging the bank's reputation.
B Major	Some criteria threatened, one of them severely affected, might endanger the success of the problem.
C Moderate	Some criteria affected but with effort can be achieved.
D Minor	Minor impact, can cause inconvenience, but it is easily rectified.
E Insignificant	Impact may be safely ignored.

Appendix B – Consequence scale

The next step involved the creation another five-point descriptive scale, this time showing the likelihoods of potential risks to occur (Appendix C)

Rating	Description
A Almost Certain	Very high probability of occurrence; can occur more than one time.
B Likely	High probability of occurrence.
C Possible	Reasonable probability of occurrence.
D Unlikely	Plausible, but unlikely to occur under normal circumstances.
E Rare	Not impossible, but very low likelihood of occurrence.

Appendix C –Likelihood scale

By using a matrix (Appendix D) to combine the consequence and likelihood ratings, initial priorities and the levels of risk were generated.

	Consequences				
	Insignificant	*Minor*	*Moderate*	*Major*	*Critical*
Likelihood	E	D	C	B	A
Almost Certain	Medium	Medium	High	High	High
Likely	Low	Medium	Medium	High	High
Possible	Low	Medium	Medium	High	High
Unlikely	Low	Low	Medium	Medium	High
Rare	Low	Low	Low	Medium	Medium

Appendix D –Priority-setting matrix

Risk Treatment

The purpose of risk treatments is to determine and identify the options for reducing the likelihood or consequences of each Medium or High risk (Cooper, et al. 2005). This process involves the identification of feasible responses to the risks, and the selection of

the best responses by determining the advantages and disadvantages of each option. Different risk management strategies can be adopted, especially risk prevention or impact mitigation strategies. The first are directed to eliminating the sources of risk or reducing the likelihood of occurrence, whereas the latter are directed to minimizing the consequences of risks (Cooper, et al. 2005).

Analysis of Potential Suppliers: Complementary Information

Criteria Formulation

One of the important aspects in the supplier selection process is the definition of criteria (Şen, et al. 2008). Many techniques and methods have been developed over the years in the field of Operations Research to support companies in dealing with the increased complexity of the supplier selection decisions (Şen, et al. 2008). An example of such methods is multi-criteria decision aid, i.e., decision taken based on multiple criteria. The process of decision based on multiple criteria rather than on a single criterion provides a more comprehensive method to assist the supplier selection process. Through a literature review, eight supplier selection criteria were regarded as important (Appendix E). These criteria consist of supplier characteristics (e.g. reputation or location) and bid/product characteristics (e.g. quality).

Bid / Product Criteria	**Relevant Aspects**
Quality	Has the product received quality awards? Are there good references?
Price	What is the price of the technology including licensing, maintenance and use costs?
Time	How long it will take to develop and implement the technology?
Flexibility	The ability to adjust volumes, i.e. is the solution scalable?
Supplier Criteria	**Relevant Aspects**
Reputation	Reputation and position in the industry.
Location	Is the supplier located in a good location for the business to go smoothly?
Support Services	How is assistance provided?

Appendix E –Description of criteria

Company A's Evaluation

Reputation: One of the biggest companies in the world in speech technology and multimodal biometrics. It has over 20 years of research, development, and implementation experience across the globe in both public and commercial sectors.

Location: Company A is headquartered in Saint-Petersburg, Russia, but has offices in the U.S.A., Mexico, Germany, and Finland.

Support Services: They are partnering with Portuguese companies. These companies are giving training to work with their technologies, in order to ensure they deliver a good level of service and reliability to the customers. However, in most cases, assistance may be provided remotely.

Quality: The company is certified according to ISO-9001:2008. The voice biometrics solution has received international awards, being distinguished both in terms of reliability and speed.

Price: The main elements of cost are the price of the licensing of the technology (licensing fee) and the number of users completing the enrollment phase (usage fee), to which may be added optional consultancy fees to help in the integration of the platform and make it scalable in the future. The licensing fee costs €18,000. To calculate the usage fee three scenarios were envisioned, one assuming little adoption, another medium adoption, and one assuming large scale adoption. The estimation of costs should not ignore internal development costs, which include the direct input of labor into the integration of the platform on the bank's existing systems architecture, overheads costs (namely the provision of financial or legal support), and contingency

costs (margin for unforeseen events). Past experience suggests that these costs may lie between €30,000 and €50,000.

Time: Solution can take some months to develop and implement, but since it is language independent, no time is needed to adapt it to Portuguese.

Flexibility: The solution is scalable but may imply contracting consulting services.

Company B's Evaluation

Reputation: Company B is a relatively small company with few years of existence (it was founded 2011). The company has little experience working with companies as it mainly engages in business with developers and provides services to end-users.

Location: It is located in the U.S.A. and has no subsidiaries or partners in Portugal.

Support Services: Support services would be provided remotely.

Quality: Even though the platform does not have support for Portuguese language, the company's assistant in the available languages has been highly rated by its users. It does not necessarily mean that the company manages to replicate the quality in Portuguese.

Price: The service offering from company B includes dedicated work on integrating their API.AI (their Application Programming Interface platform) to AB's app. The major elements of cost of this offering are: the (i) direct input of labor into the project, (ii) the cost of the technology, and (iii) the internal integration costs to which we have to add contingency costs. For the development phase four people would need to be hired: one server developer for 1 month, one Portuguese linguist for 2 months, one Portuguese writer for 2 weeks, and a Portuguese tester for 1 month. For the post-development phase, a Portuguese linguist and a Portuguese tester would need to work on the platform for 2 months (3 hours a day) to provide support. The costs of the technology include the server costs for a two-year period (€36,000), maintenance costs, and API.AI license fees

(also for a period of two years). Assuming 30,000 users[4] (i.e. number of downloads of the app), the budget for this option would be €180,000 in total (including €40,000 referring to an estimate of the internal development costs).

Time: The API.AI platform currently does not have support for Portuguese language but it is expected to have it available until the end of 2014. The development time is estimated to take up to two months.

Flexibility: The solution is scalable.

Company C's Evaluation

Reputation: Company C is a speech enabled software solutions provider with over 15 years of experience. Even though it is relatively small company (+50 employees), it has collaborated with 250 companies in different sectors (including banking) and its voice biometrics installations have reached a scale that is listed among top-5 in the world.

Location: It is a Turkey based company.

Support Services: Assistance is provided remotely as the company has no subsidiaries or partners in Portugal.

Quality: Strong investment in R&D is reflected in continuous improvement of the biometrics systems. The systems are also considered to be stable and have good performance. However, no real-life examples of voice virtual assistants were provided.

Price: The costs of both voice biometrics and voice-enabled virtual assistant for the mobile platform comprise the number of licenses (i.e. the number of downloaded mobile apps), project management costs, and an annual fee for maintenance and upgrade purposes. Even though Company C is yet to have a Portuguese voice virtual assistant, it is willing to bear with the development costs on its own.

[4] In 2014, the number of downloaded apps is ~25,000. The 30,000 is an estimate of the number of total downloaded apps in 2015.

Time: The development of capabilities for the voice assistant platform to support Portuguese takes 3-4 months, to which should be added the implementation time.

Flexibility: Both solutions are scalable.

It is important to observe that even though a first examination was already conducted, a more extensive analysis should nonetheless be done in the future. There are two reasons for this. First, in case more potential partners are contacted, they also should be scrutinized. The second reason is that, since supplier selection affects several activities within the bank, the appraisal processes should be made under the consensus of a group of people representing different departments of the bank. This group should structure the criteria into a hierarchy and define the relative weight of each of the criteria based on its importance.