MARKET STUDY OF THAI SPEECH TECHNOLOGIES IN THAILAND

Miss Puthita Wacharasin

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Engineering Management
The Regional Centre for Manufacturing Systems Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2010

การศึกษาตลาดของเทคโนโลยีเสียงพูดภาษาไทย ในประเทศไทย

นางสาว พุทธิตา วัชรสินธุ์

วิทยานิพนธ์นี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาการจัดการทางวิศวกรรม ศูนย์ระดับภูมิภาคทางวิศวกรรมระบบการผลิต
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553

Thesis Title	MARKET STUDY OF THAI SPEECH TECHNOLOGIES IN THAILAND
By	Miss Puthita Wacharasin
Field of Study	Engineering Management
Thesis Advisor	Assistant Professor Seeronk Prichanont, Ph.D.

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

..... Dean of the Faculty of Engineering
(Associate Professor Boonsom Lerdhirunwong, Dr. Ing.)

THESIS COMMITTEE

..... Chairman
(Professor Sirichan Thongprasert, Ph.D.)

..... Thesis Advisor
(Assistant Professor Seeronk Prichanont, Ph.D.)

..... Examiner
(Assistant Professor Daricha Sutivong, Ph.D.)

..... External Examiner
(James Wallbank, Dr.)

พุทธิตา วัชรสินธุ์ : การศึกษาตลาดของเทคโนโลยีเสียงพูดภาษาไทยในประเทศไทย.
(MARKET STUDY OF THAI SPEECH TECHNOLOGIES IN THAILAND)
อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร. สีรง ปรีชานนท์, 195 หน้า.

เทคโนโลยีเสียงพูด (Speech Technologies) ที่จะพิจารณาในการศึกษานี้รวมถึงเทคโนโลยีสังเคราะห์เสียงพูด (Text-to-Speech- TTS) และเทคโนโลยีรู้จำเสียงพูด (Automatic Speech Recognition- ASR) ซึ่งเป็นเทคโนโลยีใหม่สำหรับประเทศไทยและจะมีการนำมาประยุกต์ใช้กับธุรกิจในอนาคตอันใกล้ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติหรือเนคเทคได้เป็นผู้พัฒนาซอฟต์แวร์เทคโนโลยีเสียงพูดภาษาไทยและกำลังวางแผนที่จะทำการขายซอฟท์แวร์ในเชิงธุรกิจ เนื่องจากเทคโนโลยีนี้มีโอกาสเชิงธุรกิจอยู่มากดังนั้นจึงมีความสำคัญในการศึกษาตลาดเพื่อวิเคราะห์ถึงการนำไปประยุกต์ใช้ที่เหมาะสมของเทคโนโลยีเสียงพูดภาษาไทยสำหรับตลาดประเทศไทย

การศึกษานี้ใช้สองวิธีหลักในการศึกษาได้แก่ การศึกษาลักษณะการประยุกต์ใช้งานของเทคโนโลยีเสียงพูดในต่างประเทศ และการจัดการสัมภาษณ์กับกลุ่มลูกค้าที่เป็นไปได้ของเนคเทค โดยใช้เครื่องมือและเทคนิคเชิงวิเคราะห์ตลาดเพื่อวิเคราะห์หาโอกาสทางธุรกิจที่เหมาะสม จากการศึกษาพบว่าตลาดที่มีศักยภาพในปัจจุบันคือธุรกิจศูนย์บริการข้อมูลทางโทรศัพท์ (Call Center) ซึ่งจากการประมาณการมูลค่าตลาดของธุรกิจซอฟต์แวร์ศูนย์บริการข้อมูลทางโทรศัพท์พบว่ามีมูลค่าตลาดประมาณ 595 ล้านบาท หรือ 19 ล้านดอลล่าร์สหรัฐ ก็สนับสนุนว่าธุรกิจศูนย์บริการข้อมูลทางโทรศัพท์เป็นตลาดที่มีความน่าสนใจ นอกจากนั้นยังมีเหตุผลอื่นๆ เช่น ลักษณะธุรกิจมีความเกี่ยวของกับเสียงพูด ความสำเร็จของการประยุกต์ใช้ในต่างประเทศ การเติบโตของธุรกิจในประเทศไทย และอื่นๆ จากการศึกษาพบว่าลูกค้าส่วนใหญ่ยอมรับความสามารถของเทคโนโลยีสังเคราะห์เสียงพูดในปัจจุบัน แต่สำหรับเทคโนโลยีรู้จำเสียงพูดยังต้องมีการปรับปรุงเพิ่มเติม ส่วนแผนการนำเทคโนโลยีเสียงพูดไปใช้กับธุรกิจศูนย์บริการข้อมูลทางโทรศัพท์ได้แก่การนำไปใช้กับระบบตอบรับโทรศัพท์อัตโนมัติ (Interactive Voice Response- IVR) เทคโนโลยีเสียงพูดมีความเหมาะสมสำหรับงานที่ง่าย ไม่ซับซ้อน และมีการทำซ้ำบ่อยๆ ซึ่งจะช่วยลดการทำงานของพนักงานรับสายได้

สำหรับตลาดในอนาคตที่มีความเป็นไปได้ได้แก่ แอปพลิเคชันที่ช่วยให้ผู้ใช้งานเข้าถึงข้อมูลต่างๆได้ง่ายขึ้น (Accessibility Applications) และกลุ่มธุรกิจมือถือ (Mobile Applications) และเครื่องอ่านหนังสืออิเล็กทรอนิกส์ (E-Book Readers) ซึ่งจากการศึกษาทั้งหมดพบว่ามีความเป็นไปได้สำหรับการเติบโตของเทคโนโลยีเสียงพูดในประเทศไทย และในอนาคตอาจเป็นหนึ่งเทคโนโลยีสำคัญในการเพิ่มมูลค่าของผลิตภัณฑ์หรือบริการต่างๆ ได้

ศูนย์ระดับภูมิภาคทางวิศวกรรมระบบการผลิต
สาขาวิชา การจัดการทางวิศวกรรม ลายมือชื่อนิสิต.....
ปีการศึกษา 2553 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

##5271632921 : MAJOR ENGINEERING MANAGEMENT

KEYWORDS : MARKET STUDY / SPEECH TECHNOLOGIES / THAILAND

PUTHITA WACHARASIN : MARKET STUDY OF THAI SPEECH TECHNOLOGIES IN THAILAND. THESIS ADVISOR : ASST. PROF. SEERONK PRICHANONT, Ph.D., 195 pp.

Speech technologies, in this study include Text-to-Speech (TTS) and Automatic Speech Recognition (ASR), are one emerging technology in Thailand anticipated to be in use in the near future. The National Electronics and Computer Technology Center of Thailand (NECTEC) has developed Thai language speech technology software and looks forward for commercialization. With many possible business opportunities, the market study must be conducted in order to suggest the applications of Thai speech technologies that are viable in market aspect to be commercialized in Thailand's business environment.

The study was based on two main areas: the study of the applications in other countries, and conducting interviews with the potential customers of NECTEC. The study mainly uses the market analysis tools and techniques to analyze the business opportunities. From the study, the current potential market is the call center business. The result is supported by the estimated market worth of the call center software business of 595 million THB (approximately US$ 19 million), which suggests the potential market for speech technologies. Other reasons such as the related nature of the business to speech, successful international call center applications, the growth of the call center business in Thailand, and more, also support the implementation of speech technologies with the call center business. In general, current TTS capability meets the requirements of most customers, but improvement on ASR capability is required. The plan is to apply speech technologies with the existing interactive voice response (IVR) system where general tasks can be taken over by speech technologies to reduce the need of human operators for simple and repetitive tasks.

As for the future market, two interesting future markets are the accessibility applications and the mobile applications & e-book readers. It can be suggested that speech technologies have a potential to grow in Thailand's business environment, and might become one important value-added function for the existing products or services.

The Regional Centre for Manufacturing System Engineering

Field of Study : Engineering Management

Academic Year : 2010

Student's Signature Puthita Wacharasin

Advisor's Signature

Acknowledgements

I would like to thank everyone who has both directly and indirectly contributed to my master's thesis. First, I would like to thank my thesis advisor, Assistant Professor Seeronk Prichanont, who has given me support, guidance, and suggestions throughout the process. I would also like to thank all the module tutors, coordinators, and my thesis committee that gave me invaluable knowledge and suggestions required to accomplish this thesis.

Another main support for this thesis is from the National Electronics and Computer Technology Center of Thailand (NECTEC). I would like to thank Dr. Kalaya Udomvitid, the director of Business Development and Technology Transfer Division (BTT), for giving me the opportunity to work along with NECTEC on this study. Also, I would like to thank all NECTEC's staff that have contributed to my thesis and provided necessary information and data, especially Dr. Chai Wutiwiwatchai, Mr. Sawit Kasuriya, and Miss Jintana Pattanatornchai. Without the support from NECTEC, it would not be possible to accomplish this study.

Finally, I would like to thank my family for continuously supporting me and encouraging me to accomplish this thesis.

Contents

Page

Page

Page

Page

Page

Page

List of Tables

Page

List of Figures

Page

Page

Page

Page

Abbreviations

NECTEC = The National Electronics and Computer Technology Center

NSTDA = The National Science and Technology Development Agency

TTS = Text-to-Speech

ASR = Automatic Speech Recognition

IVR = Interactive Voice Response

CTI = Computer Telephony Integration

THB = Thai Baht

Chapter I

Introduction & Background

1.1 Relevant Background

1.1.1 Background of Speech Technologies

Nowadays, businesses introduce technology as a means to gain competitive advantages, enable or enhance critical business processes, and manage operations. (WMG, 2009) Speech technologies are one emerging technology in Thailand that are anticipated to be in used in the near future. Though speech technologies are not commonly known among the businesses in Thailand, the technologies have been implemented in many other countries such as the United States, the United Kingdom, and Japan for a long time already.

According to Johnston, Whittaker, and Attwater, speech technologies have been used to launch new services, add values to existing services, and differentiate the products. (Johnston, Whittaker, and Attwater, 1996) There are many types of speech technologies such as speech analysis and processing, speech compression or coding, speech enhancement, speech recognition, speech synthesis, and speaker verification, but the most interesting ones among them are speech synthesis (also known as text-to-speech technology or TTS) and speech recognition (also known as automatic speech recognition or ASR). TTS enables the conversion of text into synthesis speech that people can understand. In the opposite way, ASR converts spoken words to text or parameters that computers can understand. And this study will focus only on TTS and ASR technology.

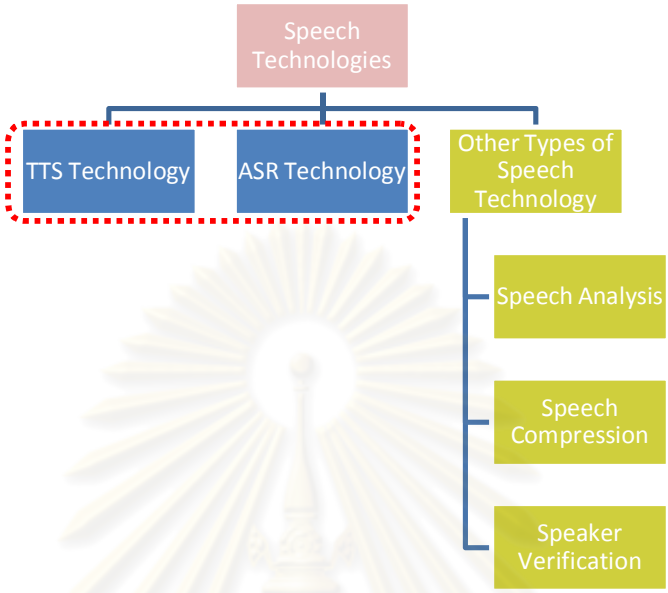


Figure 1 Speech Technologies

1.1.2 Applications of Speech Technologies in International Business Environment

In the past, despite the advanced systems demonstrated in the laboratories, not many applications of speech technologies were successful. Only the ones that allow for the limitation of technology were successful. However, Johnston, Whittaker, and Attwater suggested that the applications that are simple, cost effective, and provided a clear benefit to the user would be successful. (Johnston, Whittaker, and Attwater, 1996) Some fields that speech technology has been applied to are:

- Telecommunications: Call centers
- Desktop Applications: Voice navigation of desktop
- Mobile Phone Applications: Voice dialing / Voice command
- Automation: Voice dictation
- Transcription: Creation of reports for medical and legal field
- Game: Voice command

- Aids-to-the-Handicapped: Screen reader for blind people

As suggested above, speech technologies have been in used in the developed countries such as the United States, the United Kingdom, and Japan for a long time already. With many success stories from companies devoted to this field such as Nuance Communications Inc., Microsoft, IBM, etc. The global market for voice business technologies and services was worth US$1,105 million in 2004 with the expected annual growth rate of 30%. (Eastwood, 2005) The North America is the leading sector of the market since it is the early adopter of speech technologies. However, Asia and Pacific is becoming an emerging market of the speech technologies with a very high growth rate. Some success applications of Nuance Communications Inc. are further described in the Literature Review Section.

1.1.3 Speech Technologies in Thailand

As the national center specialized in electronics and computer technologies, NECTEC (The National Electronics and Computer Technology Center) has been devoted in research and development of speech technologies. Speech technology is one field that NECTEC has been invested to develop over ten years. There are many types of speech technologies such as speech analysis and processing, speech compression or coding, speech enhancement, speech recognition, speech synthesis, speaker recognition, and multimodal interaction.

There are only several sectors in Thailand that contribute to building the speech core technology. However, NECTEC mainly contributes to research, not much of commercialization has been done. NECTEC has developed the software for a Thai language TTS technology called VAJA 6.0 and the software for Thai language ASR technology called iSpeech. However, as an emerging technology for Thailand, the business feasibility study of the technologies must be conducted in order to identify the right business that the technologies can be applied to. The study must address the prospect

customer group of the technology and the study should be able to determine if the technology will be able to be commercialized in the business perspective.

1.1.4 Background of NECTEC

The background information of NECTEC is based on NECTEC's official website (www.nectec.or.th).

The National Electronics and Computer Technology Center (NECTEC) was established on 16 September 1986, initially started as a project under the Office of Permanent Secretary, Ministry of Science, Technology and Energy. On the 30th of October 1991, NECTEC changed its status into specialized national center under the National Science and Technology Development Agency (NSTDA), and NECTEC's organization system was also changed in order to improve its efficiency. The main responsibilities of NECTEC are:

- "To undertake, support and promote the development of electronics and computer technologies through research and development activities" (NECTEC, 2010)
- "Provides a linkage between research communities and industries through the established industrial clusters" (NECTEC, 2010)
- Funding researches of public sectors
- Support electronics, computer, telecommunication, and information industry sectors
- Act as a secretary office for NSTDA

Since 1996, NECTEC was given the opportunity from Her Royal Highness Princess Maha Chakri Sirindhorn to form a project in order to transfer knowledge to Thai people and also to implement IT system with Thai society with the emphasis on the students in the countryside, people with disability, and sick children in the hospitals.

1.1.5 VAJA 6.0: TTS System by NECTEC

VAJA 6.0 is Thai language speech synthesis software developed by NECTEC. The technology enables the conversion of Thai text into synthetic speech that people can understand. This TTS technology has been developed since 1998 and some similar technologies were developed by other sectors as well (only several sectors in Thailand are devoted to this field). However, the VAJA software has the strength that exceeds other software that it can create almost any word in Thai language because the software has the text analysis function that can analyze even the words that are not found in the dictionary. Moreover, the users can even input the new specific words such as the names and authorize their pronunciations as well. The software can be used via the serviced websites or in package software form. The software works on Linux and Windows, or serviced website. In case working as standalone, the CPU must exceed or at least a Pentium 4, and the memory must not be lower than 256 MB with the hard disk size greater than 50 MB. (NECTEC, 2010)

1.1.6 iSpeech: ASR System by NECTEC

iSpeech software is speech recognition system that can recognize the Thai language speech. The system converts spoken words into parameters that it can understand. The Thai speech recognition system started the research and development in year 2000. There are three main types of iSpeech:

- iSpeech- W: Isolated word recognition (IWR)
- iSpeech- R: Continuous speech recognition (CSR) with limited grammar
- iSpeech- N: Continuous speech recognition (CSR) with unlimited grammar

The iSpeech is easy to use and was developed in both execution file and Windows DLL. (NECTEC, 2010)

1.1.7 Applications of Speech Technologies in Thailand

Some applications of VAJA 6.0 (solutions that use only TTS technology) that have already been launched in Thailand are as follows: (NECTEC, 2010)

- NVIS : Automatic online news report is one technology that uses TTS technology of VAJA 6.0 to convert news text from daily news headline into speech. The users can call and listen to the news report from the interactive voice response system (IVR), and also the users can choose the source and type of the news. Moreover, the users can bookmark their favorite items to increase the convenience and decrease the time.
- ABDUL Help! : Automatic question answering system that enables the users to input the typed inquiry. The system will analyze the question, request the information from the database, and then answer with short synthesis speech for the users.

Some applications of VAJA 6.0 and iSpeech (solutions that combined both TTS and ASR) that have already been launched are as follows: (NECTEC, 2010)

- TVIS : Traffic voice information service combines the ASR and TTS system to report Bangkok's traffic by voice service. The users can call to the call center and say the name of the street, and the ASR system will convert the speech into text. The system will then find the information from its database and respond to the users by synthesis speech, which summarizes traffic situation on the specific street without the need of human attendants. The system can decrease the cost of hiring and training human attendants, and in the user perspective, the traffic information inform users to avoid congested areas.

1.2 Statement of Problem

The Thai language speech technology software: VAJA 6.0 (TTS technology) and iSpeech (ASR technology) have been developed by NECTEC for a long period of time. The speech technologies will be stable and will be ready to be commercialized in the near future. However, since the TTS and ASR technology are considered new technologies for Thailand, the market study must be conducted before the commercialization. The problem is that NECTEC still does not have the target market for speech technologies. At first, NECTEC planned the commercialization to target only at niche market with the applications for blind people. However, due to some policies changed, they also want to target mass market as well. Since NECTEC is Thailand's leading center that contributes to research and development of electronics and computer technology that supports industries in Thailand, the technology must be able to be implemented and commercialized.

One market that NECTEC has recently been interested in is the call center business. Since the call center has become increasingly important part of almost all types of business and is directly related to the increase in customers' satisfaction, the technology improvement can be one way to increase the service quality. The TTS technology can be implemented along with the ASR system. TTS technology enables the computers to read text to the customers such as listening to traffic report, whereas ASR can receive simple data entry from the customers such as entering the street names by voice. Currently, most call centers in Thailand implement interactive voice response (IVR) system. IVR system enables the self-service system for customers either by keyboard input or speech recognition. In Thailand, most IVR system uses only keyboard input and responds with pre-recorded audio, but in reality some data cannot be input by the keyboard such as the address, and some data cannot be pre-recorded. According to the international models from other countries, call center is among the businesses that most widely-used speech technologies. The detailed market study must be conducted to ensure the right market for speech technologies to be commercialized.

1.3 Objective of Thesis

The objective of the thesis is to suggest the applications of Thai speech technologies (TTS and ASR) that are viable in market aspect to be commercialized in Thailand's business environment. The thesis aims to:

1. Identify the potential market of the applications of existing speech technologies in Thailand.
2. Study the future potential market for further technology development.

1.4 Scope of Thesis

The study will be conducted in two main areas: market study for existing technologies (technologies at current level) and market study of future potential market

Market Study for Existing Technology:

- Assessment of existing technologies' capabilities
- Study of successful applications in other countries
- Market research by interviews
- Identification and segmentation of possible markets
- Targeting of the most appropriate segment
- Positioning of the technologies
- Approximate the market worth

Market Study of Future Potential Market

- Suggestion of future potential market of speech technologies
- Suggestion of future capabilities needed

1.5 Expected Benefits

From this study, the main benefit of the study is that it can be a decision making tool for the future commercialization that will benefit both NECTEC (the developer of speech technologies) and its customers (the businesses that use the speech technologies to develop applications such as software houses).

The expected benefits are:

1. Identified potential market group of current speech technologies in Thailand
2. Estimated market worth of current potential market group
3. Application plan of the speech technologies for the target group
4. Identified potential market group(s) of future speech technologies in Thailand
5. Suggestion of future capabilities needed in order to target both the current and the future potential market group(s)

1.6 Methodology

The study starts with gathering related information and studies the related theories about market study, speech technologies and their applications in other countries, the status of the technologies in Thailand, and the potential application areas.

The information is gathered from both primary and secondary source. The information can be gathered from the primary source such as phone interviews and deep interviews. The preliminary study may be conducted by phone interviews with software houses in many different businesses, which are the 1st-tier customers of NECTEC. The software houses use the speech technologies to develop into applications then sell them to their customers in different industries, which are the 2nd-tier customers of NECTEC. For example, the software houses that provide call center solutions will have customers in

different industries such as banking, manufacturing, travel & leisure, etc. It can be summarized in the diagram as follows:

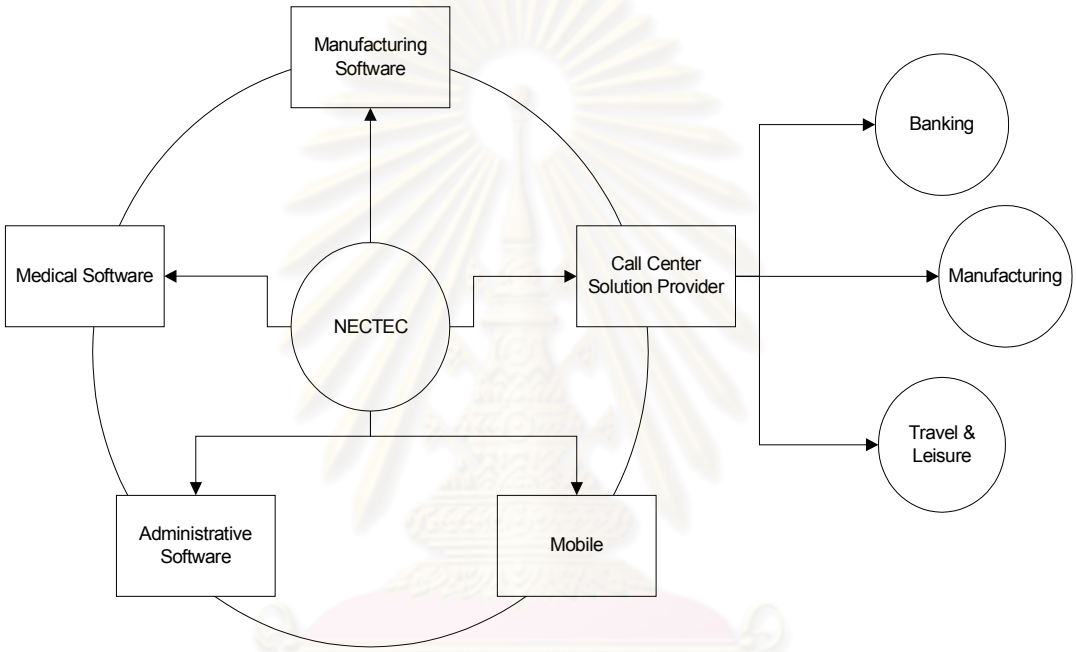


Figure 2 Value Chain of Speech Technologies

Phone interviews and some deep interviews with software houses in different businesses will be conducted. The preliminary interviews, the current capability analysis of speech technologies, and the study of applications in other countries will be used to conduct analysis in terms of marketing that will suggest the type of business that is most prospective to commercialize speech technologies. After that, the deep interviews with the selected business and its customers (2^{nd}-tier customers of NECTEC) will be conducted. The activities in each process are further explained in the “Thesis Schedule & Activities in Each Process” section.

1.7 Thesis Schedule & Activities in Each Step

	2010			2011	
	October	November	December	January	February
1. Identify problem and gather relevant information					
2. Literature review					
3. Market Study for Existing Technologies					
4. Market Study of Future Potential Customers					
5. Conclusion and thesis write up					

Table 1 Thesis Schedule

1. Identify problem and gather relevant information

The problem of the thesis is identified as stated above. After the problem is identified, related information about the status and business environment of speech technologies in Thailand and the speech technology software by NECTEC (VAJA 6.0 and iSpeech).

2. Literature review

The study of the literature should include market study, the basic mechanism behind speech technologies (TTS and ASR), the speech technology applications in other countries, and the potential applications launched by NECTEC. The tools that are relevant such as SWOT analysis should also be investigated.

3. Market Study for Existing Technologies

The preliminary market information can be gathered by conducting market research by phone interviews with the contact companies of NECTEC (software houses) to analyze the market situation in Thailand. Secondary information should also be gathered as well. The study of international applications and the analysis of capabilities of Thai speech technology software along with the interviews will be used to determine the most appropriate customer group(s). After that, the deep interviews with selected business and its customers (2^{nd}-tier customer of NECTEC) will be conducted. Data and information are analyzed and summarized.

4. Market Study of Future Potential Customers

The study of successful international applications will be conducted to suggest the future application(s) and its potential market(s) for NECTEC in order to be a guideline for the research and development of speech technologies in the future. Since some customer group(s) might be the prospect group to use speech technologies according to the international applications, but the capabilities of existing technologies cannot fulfill those needs. More interviews might be conducted to find the feedback of the requirements that the future potential groups require in order to purchase the technologies.

5. Conclusion and thesis write up

Finally, all the aspects are to be summarized and come up with the conclusion. Suggestions and further analysis of the thesis may be proposed. The last step is the thesis write-up and preparation of thesis examination.

Chapter II

Literature Review

2.1 Market Study

The market study concept in this study is based on Clifton and Fyffe's book called "Project Feasibility Study Analysis: A Guide to Profitable New Ventures".

According to Clifton and Fyffe, market study involves the identification, isolation, description, and quantification of the market. The brief description of the market, the analysis of the past and present demand, the estimate of the future demand, the positions of the product, and the estimate of the project's share of the market are involved in the study. (Clifton and Fyffe, 1977)

The market study starts with defining the objectives of the study, and Clifton and Fyffe stated that the overall objective of the market study is "to measure and forecast the market to determine if the project will product the right product at the right time and at the right price." (Clifton and Fyffe, 1977) Then the situational analysis of market is conducted by searching for secondary and primary data. The primary data can be obtained by designing questionnaire and collect sample. After the data are collected and analyzed, the total market size is to be identified and also the market share of the project is identified. The study is continued by the demand forecasting and the further suggestion of marketing plan may be suggested. (Clifton and Fyffe, 1977)

2.2 SWOT Analysis

SWOT analysis is used to define the strengths, weaknesses, opportunities, and threats of either a company or it can be broken down into smaller units such as products. In this case the SWOT analysis of Thai speech technologies is analyzed. According to O. C. Ferrell and Michael D. Hartline, SWOT analysis is one of the most effective tools to analyze market data and information. It provides the assessment of a product's capabilities

(strengths and weaknesses) against the external business environment (opportunities and threats). (O. C. Ferrell and Michael D. Hartline, 2008) Some issues to be considered in SWOT analysis are as follows:

EXHIBIT 5.5 POTENTIAL ISSUES TO CONSIDER IN A SWOT ANALYSIS

Potential Internal Strengths

- Abundant financial resources
- Well-known brand name
- Number 1 ranking in the industry
- Economies of scale
- Proprietary technology
- Patented processes
- Lower costs (raw materials or processes)
- Respected company/product/brand image
- Superior management talent
- Better marketing skills
- Superior product quality
- Alliances with other firms
- Good distribution skills
- Committed employees

Potential External Opportunities

- Rapid market growth
- Complacent rival firms
- Changing customer needs/tastes
- Opening of foreign markets
- Mishap of a rival firm
- New product discoveries
- Economic boom
- Government deregulation
- New technology
- Demographic shifts
- Other firms seeking alliances
- High brand switching
- Sales decline for a substitute product
- Changing distribution methods

Potential Internal Weaknesses

- Lack of strategic direction
- Limited financial resources
- Weak spending on research and development
- Very narrow product line
- Limited distribution
- Higher costs (raw materials or processes)
- Out-of-date products or technology
- Internal operating problems
- Internal political problems
- Weak market image
- Poor marketing skills
- Alliances with weak firms
- Limited management skills
- Undertrained employees

Potential External Threats

- Entry of foreign competitors
- Introduction of new substitute products
- Product life cycle in decline
- Changing customer needs/tastes
- Declining consumer confidence
- Rival firms adopting new strategies
- Increased government regulation
- Economic downturn
- Change in Federal Reserve policy
- New technology
- Demographic shifts
- Foreign trade barriers
- Poor performance of ally firm
- International political turmoil
- Weakening currency exchange rates

Figure 3 Potential Issues to Consider in a SWOT Analysis

(O.C. Ferrell and Michael D. Hartline, 2008)

2.3 Porter's Five Forces Analysis

Michael E. Porter has suggested five forces to analyze an industry's attractiveness in the book "On Competition". According to Michael E. Porter, Porter's Five Forces Analysis is a framework to analyze an industry's attractiveness and its profitability by understanding the underlying five forces. The five forces suggested by Porter are: customers, suppliers, industry rivals, potential entrants, and substitute products. By analyzing these forces, the industry's structure and the nature of the competition of the industry are defined. Even though different industries might have different natures, the underlying drivers are the same. If the forces are intense, the profitability of the industry might not be attractive. (Porter, 2008)

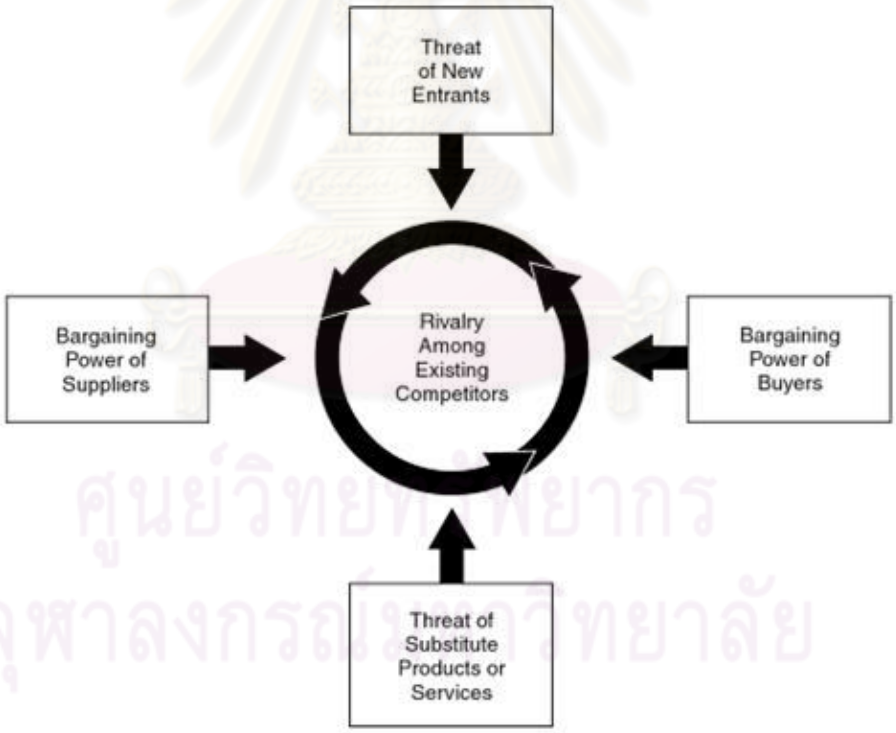


Figure 1.1 The Five Forces That Shape Industry Competition

Figure 4 Porter's Five Forces (Porter, 2008)

Threats of Entry

New entrants to the industry can increase pressure on the price and cost. Competition can increase when new entrants enter an industry. The threat of entry is related to the profit potential of an industry. If the threat is high, the existing players in the industry must keep their prices stable and keep the quality level up to standard and even increasing. The threat of entry depends on the entry barriers. If the entry barriers are low, the threat of entry is high and the profit potential is low. Other than the cost side, the assessment in terms of how the existing players in the industry will react must be carried out. If the reaction is strong, the profit potential can be low. (Porter, 2008)

The Power of Suppliers

The power of the suppliers can have high impact on industry such that powerful suppliers can lead to the higher price, limit quality, and cost shifting to the industry players. Powerful suppliers can lead to the limited profit potential because the high costs that are not covered. (Porter, 2008)

The Power of Buyers

The powerful buyers can force down prices or demand higher level of quality and service, which leads to the lower profit potential. The power of buyers can be assessed in two main aspects: bargaining leverage and price sensitivity. (Porter, 2008)

The Threats of Substitutes

The substitute products can perform the same or similar function to the existing product. If the threat of substitutes is high, there can be impact on the ceiling price, and therefore leads to the lower profit potential. (Porter, 2008)

Rivalry Among Existing Competitors

The rivalry among existing competitors directly affects the profit potential. If the rivalry is high, the profit potential tends to be low. Some examples of the rivalry are price discount, new product introduction, and advertisement. The rivalry can be assessed in two main aspects: the intensity of the competition and how they compete. (Porter, 2008)

Not only the intensity of competition that determines the rivalry, but also how companies compete is important. The price-related competition creates the greatest impact for the industry. If the companies compete by price cut, the result is destructive for the whole industry. (Porter, 2008)

2.4 Speech Synthesis (also known as text-to-speech technology or TTS)

2.4.1 Definition of Speech Synthesis

Speech synthesis technology enables the conversion of text into synthetic speech that people can understand. Thierry Dutoit stated that “Intuitively, the ultimate goal of a text-to-speech (TTS) synthesizer is to read *any* text, whether it was directly introduced in the computer by an operator or scanned and submitted to an optical character recognition (OCR) system. Reading should be intelligible and natural”. (Dutoit, 1997) He also stressed that the TTS system is different to the pre-recorded voice that the TTS system produces the *new* sentences automatically.

2.4.2 Mechanism of Speech Synthesis

According to NECTEC, the model of TTS system is presented below. There are three main functions to the system: text analysis, prosody, analysis, and speech synthesis. The text analysis function analyzes the input text and breaks the text into small part of speech called “phoneme”. The prosody analysis function analyzes the “prosody”, which are the pitch and the rhythm of the speech. The speech synthesis function takes the transcription of the text analysis and prosody analysis function and creates speech waveform. There are

many technical ways to synthesize speech; in general, there are three main ways: (NECTEC, 2010)

1. Formant Synthesis
2. Articulation Synthesis
3. Concatenation Synthesis

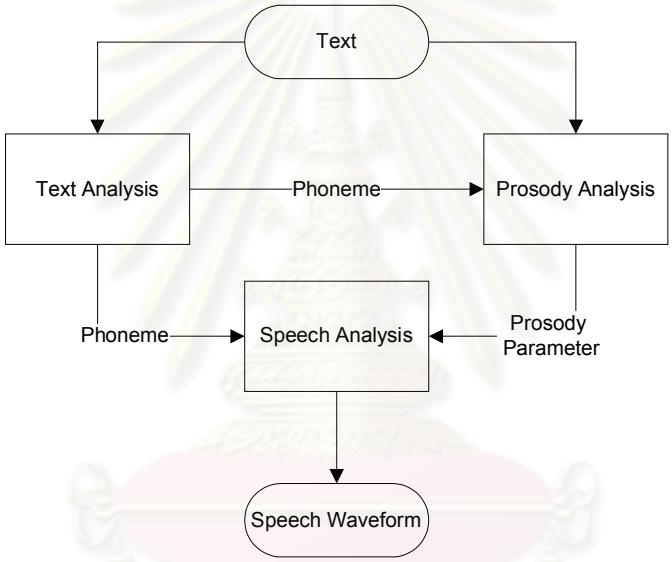


Figure 5 Model of Text-to-Speech (TTS) System (NECTEC, 2010)

2.4.3 Applications of Speech Synthesis

Many applications can be developed from this technology such as the applications for blind people (screen reader), applications for normal people (electronic mail reader using telephone interface), applications for learning objectives such as talking dictionary, and applications in telecommunication field such as call center.

2.4.4 Advantages and Disadvantages of Speech Synthesis

Advantages:

- Easy to use
- Can be performed along with other tasks (hands-free and eyes-free)

Disadvantages:

- Limited grammar and vocabulary
- Limited expressions of feelings

2.5 Speech Recognition (also known as automatic speech recognition or ASR)

2.5.1 Definition of Speech Recognition

There are many definitions for speech recognition; in general, speech recognition technology converts spoken words to text or parameters that computers can understand. Young gave the definition of speech recognition as follows: “Speech recognition systems generally assume that the speech signal is a realisation of some message encoded as a sequence of one or more symbols.” (Young, 2006)

2.5.2 Mechanism of Speech Recognition

According to NECTEC, the spoken speech spoken by human is the “speech waveform” or a sequence of symbols. However, the computer cannot directly understand the language spoken by human; therefore, the computer converts the human’s speech into parameters that it can understand known as “speech vectors”. (NECTEC, 2010)

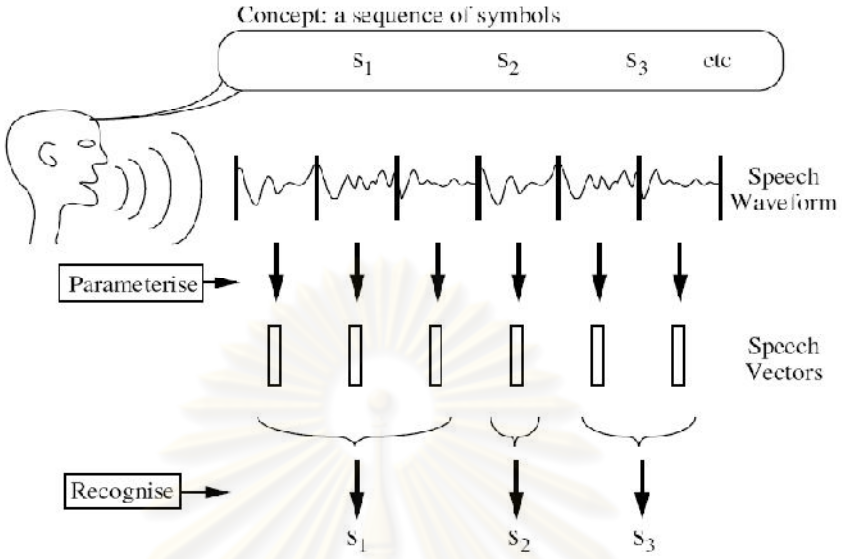


Figure 6 Message Encoding/Decoding (Young, 2006)

According to NECTEC, the model of ASR system is presented below. The ASR system receives the speech utterance as the input. The speech is extracted into small parts and sent to speech engine. Using acoustic models, the system is able to define the word match from the sound. Then the sentence match is made according to the language models, in which the words are defined by grammar rules. (NECTEC, 2010)

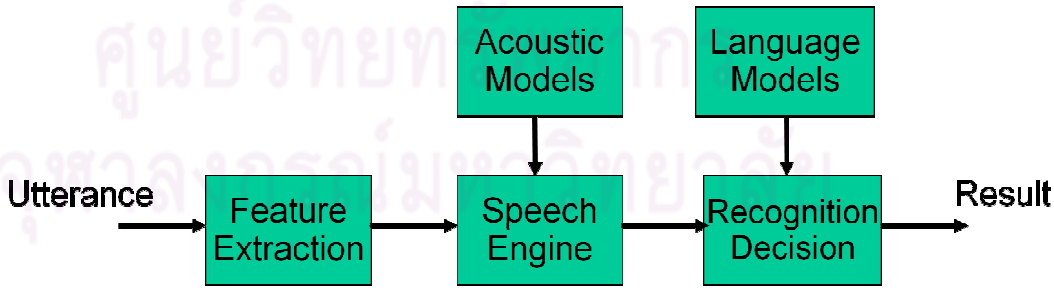


Figure 7 Modern Model of Automatic Speech Recognition (ASR) System (NECTEC, 2010)

According to NECTEC, the ASR system can be classified by four main ways, and each are divided into subsets as follows: (NECTEC, 2010)

1. Continuity of speech
    - Isolated Word Recognition (IWR)
    - Continuous Speech Recognition (CSR)
2. Speaking style
    - Isolated words or phrases – *voice command*
    - Connected speech – *digit string*
    - Read speech – *dictation*
    - Fluent speech – *broadcast news*
    - Spontaneous speech – *conversation*
3. Speaker dependency
    - Speaker dependent
    - Speaker independent
4. Unit of reference template
    - Word unit
    - Subword unit
        - Phoneme

2.5.3 Challenges for the Speech Recognition

According to Detlev Arter, some challenges for ASR system that the technology providers are trying to overcome are: (Artelt, 2008)

- Speaker variance- the extreme range of voices and individual characteristics.

- Ambient noise- the environment noise can interrupt the recognition.
- Transmission- the quality of voice transmission over the channels such as telephone lines.

2.5.4 Applications of Speech Recognition

Some applications of speech recognition are voice user interfaces such as voice dialing for smart phones, call routing, voice commanding for disables, simple data entry such as entering a credit card number, speech-to-text processing such as word processors or emails, and direct voice input in aircraft.

2.5.5 Advantages and Disadvantages of Speech Recognition

Advantages:

- Easy to use
- Fast and convenient
- Can be performed along with other tasks (hand-free and eye-free)

Disadvantages:

- Accuracy issue of the spoken language (the clearness of the speaker's voice)
- The surrounding noise can interrupt
- The customers' willingness to try new technology (as it has rarely been implemented in Thailand)
- Create annoyance for other people in some situations
- Cannot be used with confidential information

2.6 Interactive Voice Response (IVR)

Interactive Voice Response (IVR) uses automated voice to response to the customers. A spoken dialogue between human and computer can be performed by using TTS and ASR system. IVR system enables the self-service system for customers either by keyboard input or speech recognition. In Thailand, most IVR system uses only keyboard input and responds with pre-recorded audio, but in reality some data cannot be input by the keyboard such as the address, and some data cannot be pre-recorded.

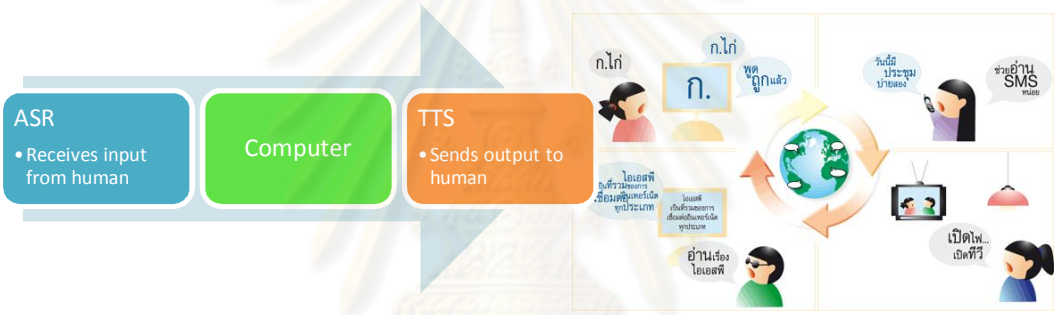


Figure 8 Interactive Voice Response (IVR) System (NECTEC. 2010)

2.7 Value Chain of Speech Technologies

According to Detlev Artelt, voice applications consist of three basic parts: voice hardware and software and voice services. The choices for companies are to buy their own hardware and software or to lease them from a hosting provider. (Artelt, 2008) The value chain of speech technologies is summarized in the following figure and each part is described according to Detlev Artelt in details in the following paragraphs:

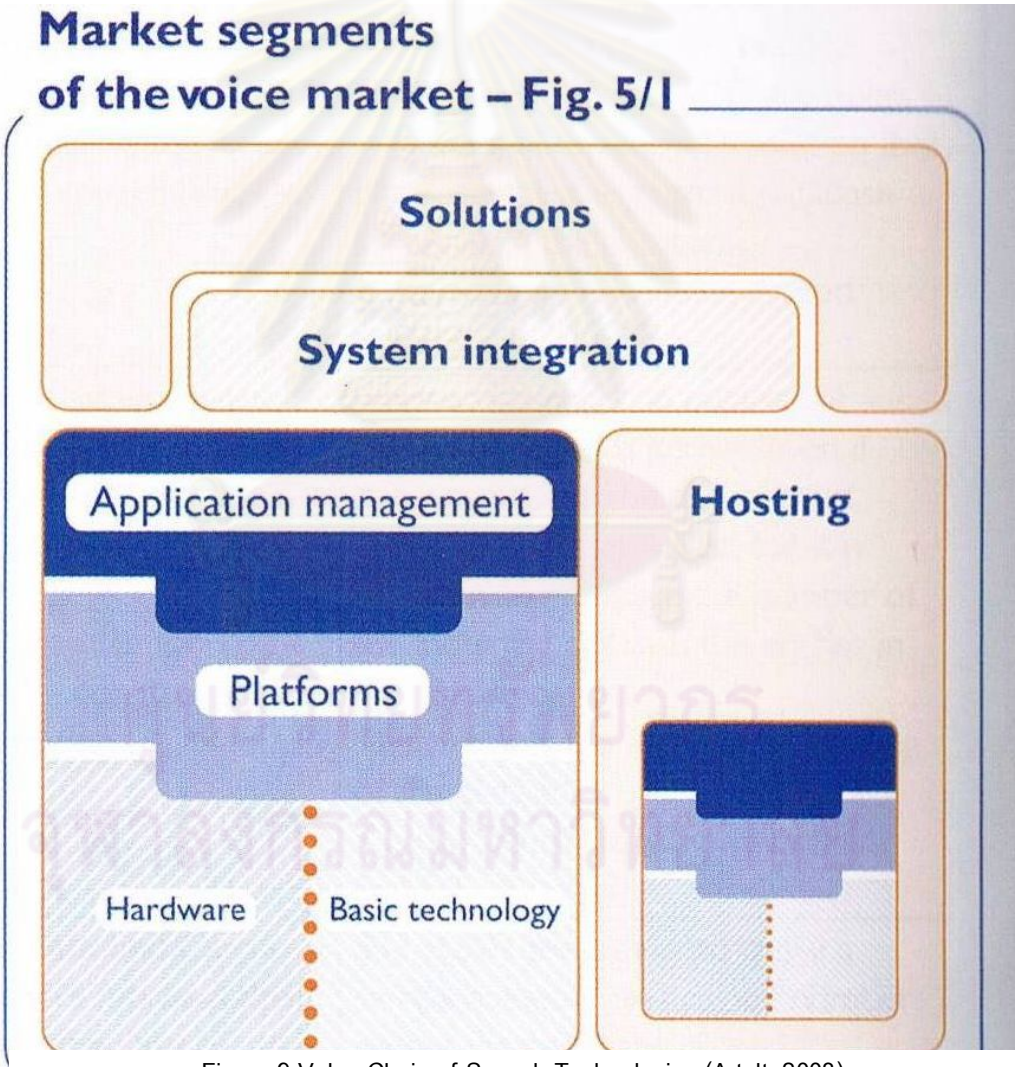


Figure 9 Value Chain of Speech Technologies (Artelt, 2008)

From "Voice Compass: International 2008/2009: Speech Goes Mainstream"

Hardware

The most essential hardware required for speech applications is a computer. In the companies, the servers are required since there is more than one computer used. Other than the computers, the interface between computer and phone network is also required.

Basic Technology

The basic technology is the mechanism that allows the computer to understand the words and synthesis the words, in this case is the TTS and ASR software.

Platforms

The platform integrates individual parts to work together as a whole. In this case, the platform is the operating system of the computer or server that allows the speech technology software to operate on. Examples of platforms are Windows, Open Source (Linux), UNIX, and mobile platform.

Application Management

The basic technology and platform are integrated into the applications for the users. The user interface is created to manage the flow of interactive discussion between the user and the computer system. The developers must decide which application they need to develop and the flow of the application.

Hosting

Hosting is to rent the system for the technologies. There are two types of hosting: server hosting and server housing. Server hosting means a company rents the hardware and the software from the provider, and the company pays only for what they use. Server housing means you rent only the facilities (hardware) and you use your own software. This is a good idea in case the companies do not want to handle with complex technologies themselves.

System Integration

The system integration integrates the applications with the existing IT systems such as databases or telephone systems. The integration between the computer and telephone is called “computer-telephony integration” or CTI, which is one most important system required in the call center system.

Solutions

All the components are combined into a solution. The solution provider provides value to the customers by delivering ready-to-use solutions. The main advantages for a company to buy a finished solution are the short lead time and the low cost. However, the disadvantage is the company is becoming dependent on the provider, and in case the company needs an update or change, there can be costs added to it.

Chapter III
Study of International Applications

3.1 Status and Capabilities of Speech Technologies in Other Countries

3.1.1 Speech Technologies Development & Capabilities

Beginning of 1990s

According to Detlev Artelt, the real speech applications came into market existence in the beginning of the 1990s. In the start, the ASR system was far from what present ASR system could do, only limited amounts of words are understood by the system. The limitations in the past were not only the result of the undeveloped speech technologies, but also due to the limitations of the computer systems and database capacity. (Artelt, 2008)

Beginning of the New Millennium

Then around the end of the millennium, many investors started to push the speech applications into the market. And some commercial applications were launched during that time. The basic technology was considered sophisticated at that time.

Present

According to Detlev Artelt, currently due to the increasingly higher computer performance and sophisticated basic speech technology, the speech applications are replacing the old touchtone system. The interactive voice response system (IVR) started off with the touchtone system, which required the users to enter numbers by using keypads. There are limitations to the touchtone system because some data cannot be entered via keypads such as names. Therefore, speech becomes a better option to communicate with the systems. (Artelt, 2008)

Base on Artelt's book, the capability of current state-of-the-art ASR system is able to understand complete continuous sentences and also extract information from them. The

users communicate with the system in “dialog” similar to human-to-human conversation. As for TTS system, the technological advancement in the past few years made the naturalness of the speech increased significantly that it is very difficult to tell the difference between the speech synthesizer and a human. In the past few years the techniques of speech technologies have advanced significantly. The innovations in speech applications are constantly increasing, mainly in the telephone-based applications, and also other areas as well. (Artelt, 2008)

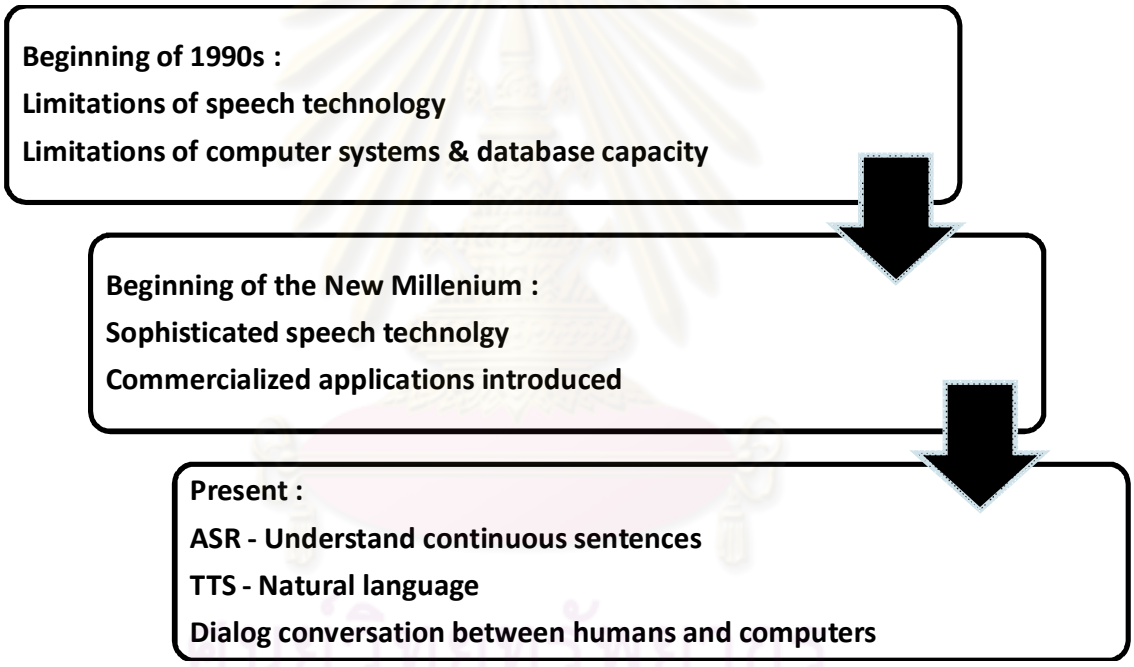


Figure 10 Speech Technologies Development & Capabilities

3.1.2 State-of-the-Art ASR and TTS Technology by Nuance Communications

According to Nuance Communications, Nuance recognizer is designed to provide the industry’s state-of-the-art ASR engine. The latest version of Nuance recognizer is claimed to have the reduced error rate compare to its previous version such as

OpenSpeech® Recognizer 3 and Nuance 8.5. The reductions of error rates compare to the previous version are as follows:

Language	Achieved RERR% vs. OSR3	Achieved RERR% vs. Nuance 8.5
U.S. English	27%	26%
Australian English	35%	29%
UK English	15%	32%
German	33%	16%
Canadian French	27%	39%
French	14%	N/A
Spanish	45%	N/A
Indian English	27%	N/A

Table 2 Relative Error Rate Reduction (RERR) for Nuance Recognizer (Nuance Communications, 2010)

The Nuance recognizer effectively ignores background noises and increases the accuracy rate for wireless and hand-free situations. This allows the users to use at any environment such as outdoor environment. The latest version also allows the differences in dialects and accents, which improves the accuracy rate of the recognizer. (Nuance Communications, 2010)

3.1.3 Current Status of Speech Technologies

According to Detlev Artelt, the users of speech applications such as telephone users have accepted the speech applications. The benefits from employing speech applications are there for both the customers and the providers. The customers benefit from the convenience of using voice with simple dialogs rather than pressing many keypad entries. Also, some main vendors started to embed built-in speech applications such as in Windows. Therefore, speech started to become more common with today's technology. (Artelt, 2008)

According to Global Industry Analysts on the “Speech Technology: A Global Strategic Business Report”, speech technologies are seen as a medium to long term cost-saving tool. Companies are becoming more interested in speech technologies due to the recession period, which encouraged them to plan the cost-saving strategy for the future. Surely the speech technologies will enable the cost-saving by reducing the labor cost and also increase the efficiency of a system. An example given by the Global Industry Analysts, Inc. is the ASR system in call center that reduced 50% of the agent cost and also improve the overall productivity. (Global Industry Analysts, 2010)

Detlev Artelt also suggested that the prices of the individual components of the speech applications will go down due to the increasing production and increasing competition. Together with the advancement in computer technology such as increasing processing power, this will eventually leads to the increasing applications in new market. (Artelt, 2008)

3.1.4 Numbers from Market Research

According to Global Industry Analysts, the forecasted global speech technology market is expected to reach US$20.9 billion by the year 2015. The global demand for ASR technology is expected to grow at a double-digit growth rate during 2007 to 2015 due to the change in lifestyle of people. These days mobility is becoming an increasingly important issue, people tend to have less time and tend to do things simultaneously. ASR technology allows user to use things hand-free and eye-free, which is suitable for busy lifestyle. An upcoming trend of ASR is the mobile applications and applications for automobiles. (Global Industry Analysts, 2010)

The current growth of global speech technology market is lower compares to the growth before the world’s recession. However, the companies’ plan of employing speech technologies prior to the recession will keep the growth rate growing. Also the companies plan to use speech technologies to decrease cost and add new channel for e-commerce as

well. Currently, the revenues for TTS in the United States are expected to grow at the compound annual growth rate (CGAR) of 23.2%. (Global Industry Analysts, 2010)

As for Europe, the analysis from Datamonitor suggested that the growth for Europe market growth is approximately 10%. The main markets are the United Kingdom, France, and Germany, which made up two-thirds of the European market. (Datamonitor) The market growth in Europe is approximated in the following figure:

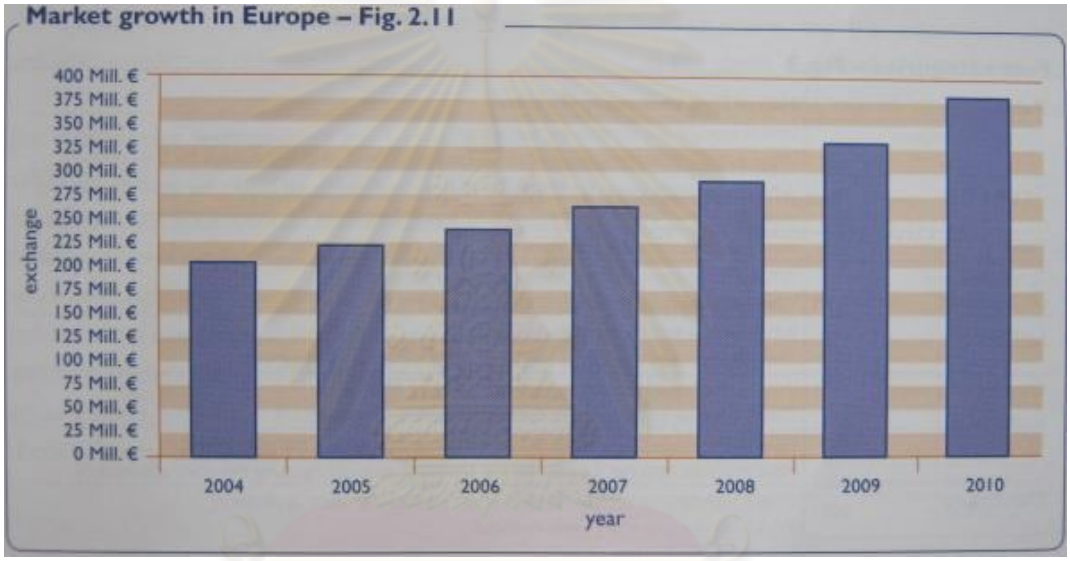


Figure 11 Market Growth in Europe (Artelt, 2008)

From "Voice Compass: International 2008/2009: Speech Goes Mainstream" (2008)

3.1.5 Key Players in the Speech Market

According to Global Industry Analysts, the some key players in the speech market are:

- Acapela Group
- Applied Voice & Speech Technologies Inc
- Convergys Corp

- Fonix Speech Inc
- Genesta Partnership
- Genesys Telecommunications Laboratories Inc
- HomeSeer Technologies LLC
- IBM Voice Systems
- Intel Corporation
- iVoice Technology
- Microsoft Speech Technologies
- Nortel Networks Corporation
- Novauris
- Nuance Communications
- Sensory Inc
- Speech + Mobility
- Syntellect
- Telephonetics VIP Ltd
- Telisma
- Vocollect Inc
- Voxify Inc
- Voxware Inc
- Wizzard Software Corporation

3.2 Successful Applications in Developed Countries

This part of the study will include the description of speech applications that might be applied as prospective applications for Thailand's business environment. The selected applications are suggested in the book "Voice Compass: International 2008/2009: Speech

Goes Mainstream" and the applications of Nuance Communication (www.nuance.com). Nuance Communications Ltd. is one of the largest speech technology solution providers in the world that uses voice to help people live and work more intelligently. With their well-known capability in providing speech technology solutions for businesses, their applications might be applied for Thailand's speech applications. The speech applications are segmented by the solution types.

Figure 12 Nuance Logo (Nuance Communications, 2010)

Accessibility

The accessibility applications increase accessibility among people by using TTS and ASR technology. It can enable hands-free work when creating documents, accessing data, or navigating their computer desktop for people with disability or injuries that prevent them from typing. Some accessibility applications are: (Nuance Communications, 2010)

Application	Technology Used	Details
Dictation by speech (create text from speech)	ASR	Employees with musculoskeletal disorders (MSDs), which include repetitive strain injuries (RSIs), can use their speech to input text and data by their voice. Talking instead of typing can

		reduce their injuries and pain, and also talking is faster than typing.
Navigating computer desktop	ASR	Both normal employees and handicapped employees can use this function to navigate their desktop by voice rather than mouse clicks.
Help people with speech impairment to learn	ASR+TTS	Blind and deaf people can use voice as a medium to learn.
Screen reader for blind people	TTS	Blind people can now read the text on the computer screen by listening to the text.

Table 3 Accessibility Applications (Nuance Communications, 2010)

Automation by Voice

Voice can be used as a remote control in situations that the users are not hands-free. An example of using voice to automate is the house appliance control: (Artelt, 2008)

Application	Technology Used	Details
Controlling media devices by voice	ASR+TTS	Voice remote control for media devices such as television, video, radio, etc. can be created. ASR can be used to input the command, and TTS can be used to respond to the user's command. For example, the video programs can be searched easily by voice. This application can be applied for

		office use with business purpose such as presentations in conferences.

Table 4 Automation by Voice Applications (Artelt, 2008)

Automotive

In automotive applications of speech technologies, the speech is used as the method of interaction to control and command system and navigation functions hands-free. The objective is to minimize driver distractions for safety reason when driving. The drive came from the lifestyle of people these days that tend to spend longer time in their cars. Therefore, there is demand from car users that they want to be connected to the world hands-free and eyes-free. One successful example of automotive solution launched by Nuance Communications is Ford SYNC, the in-car communications and entertainment system offered in most Ford, Lincoln, and Mercury models. Some automotive applications are: (Nuance Communications, 2010)

Application	Technology Used	Details
Navigation system	ASR+TTS	Enter and search for address by voice for navigation system, and get a response from the system by read-out route guidance and traffic alerts. A navigation system requires a lot of input from the users, and voice can be the convenient solution to use rather than using a pen or buttons.
Mobile communication	ASR+TTS	The in-car communication by voice such as phone calls can be initiated by voice. The users can dial

		calls by saying the names or the numbers, and also they can pick up calls by voice. The users can also listen to text messages read out loud by TTS technology.
Control in-vehicle features	ASR+TTS	The car users can control features such as music tuning by voice. The users can select the songs by the title, artist, album, or playlist, or select the radio station by voice. The system is able to respond to the selected music by TTS technology.
Real-time traffic information and services	ASR+TTS	The users can access information only by using their voice. The database is connected with the vehicle and can be easily accessed by voice. Some available services are such as real-time weather and traffic information.

Table 5 Automotive Applications (Nuance Communications, 2010)

Figure 13 Figure Ford In-Vehicle System

From Ford Motor Company

Call Center

Researches show that phone is still the most heavily used customer service channel; it becomes a business requirement to delivery good phone services. Sometimes a caller needs to hear dynamic information that cannot be pre-recorded such as names, addresses or information from a database. TTS system can read back the dynamic information as part of an automated system, without transferring to an agent to complete the task. As for ASR system, it can handle calls that cannot be handled with touchtone system such as collecting names or address, or it can be used to avoid long lists of options.

Detlev Artelt suggested that the speech technologies can be implemented in combination with human operators in call centers. The speech applications can take over the general inquiries, which are simple and repetitive, while the human operators can support the speech applications in tasks that are more complex. By implementing speech technologies does not mean the human operators are made redundant, but rather giving them more time for more complex tasks. (Artelt, 2008)

Some applications for call center are: (Nuance Communications, 2010; Artelt, 2008)

Application	Technology Used	Details
Authenticate customers	Speaker Verification	Speaker verification can be used to authenticate customers instead of using password.
Call routing	ASR+TTS	Call routing can be done using ASR technology. The caller can request by voice for the system to call the directory requested.
Menu by voice	ASR+TTS	The menu directory can be implemented by voice in combination with touchtone system. Some long-listed menu or menu that cannot be selected by keypads

		may be implemented by ASR. An example would be the movie hotlines to check the movie schedules, movie titles and time can be input by voice.
Answer general inquiries	ASR+TTS	General inquiries can be answered by ASR and TTS. Inquiries such as account balance or contact numbers of bank branches can be answered by speech applications rather than human operators.
Ordering system	ASR+TTS	Ordering system by voice can be very suitable due to its nature. The necessary data for the order system are item numbers, amount, size, and color. These data are more suitable for voice rather than keypads. By implementing speech applications, the order system can work 24/7, which will fully respond to customers' demand.
Collect names and addresses	ASR+TTS	Names and addresses which cannot be input by keypads can be input by voice.
Read data that cannot be pre-recorded	TTS	Some data such as names or dynamic information that changes frequently cannot be pre-recorded, can be read by TTS technology.
Outbound notifications	TTS	Outbound notifications can be created easily by TTS technology without the need of human operators.

Table 6 Call Center Applications (Nuance Communications, 2010; Artelt, 2008)

The speech applications can be applied to call centers in all verticals such as financial & banking, health care, government sectors, travel & tourism, telecommunications, logistics, and all types of private enterprises that require call centers.

Mobile Applications

Mobile applications are another field that speech can be applied to due to the nature that supports the speech applications. The text or message can be created by voice, as well as the voice dictation for emails is enabled. The online search from websites such as google can be done by voice. Also, the command and control functions of the mobile phones can be handled via ASR technology. Some applications for mobiles are: (Nuance Communications, 2010; Artelt, 2008)

Application	Technology Used	Details
Create text messages by voice	ASR	Instead of typing text messages, text messages can be created by voice dictation.
Read text messages by voice	TTS	Text messages can be read out loud by TTS technology without requiring eyesight from users.
Voice dictation for emails	ASR+TTS	Emails can be created and read by voice applications.
Online search	ASR	Online search from websites such as googles are provided in mobile platform, which enable on-the-go search of information.

Voice command for mobiles	ASR	Functions in mobile phones can be controlled by voice without the need of keypads such as voice dialing.

Table 7 Mobile Applications (Nuance Communications, 2010; Artelt, 2008)

Employee Productivity

The employee productivity suit is used in companies to increase a company's workforce productivity. The employee productivity suit can be used to ensure consistent in business processes and employee communications. By simplifying business processes, the IT costs can be decreased as well. The solution provides employee directory assistance and self-service applications that increase workforce productivity, which reduce IT costs. Some applications for the employee productivity suit are: (Nuance Communications, 2010)

Application	Technology Used	Details
Internal dialer / Automated attendant	ASR	Internal dialer can eliminate the need of remembering the internal numbers by routing calls by voice. The ASR technology allows the call routing to the specific person in the organizations without the need of human operators, which increase the productivity and decrease the waiting time for the callers.
Password reset function	Speaker Verification	Companies face repetitive task of password resetting for the employees. Password reset using voice biometric technology can be an alternative way to decrease IT costs. Gartner Group suggested that approximately 30% of all calls to the IT department

		are related to password. (Gartner Group, cited in www.nuance.com) Employees can use their voice to identify themselves to reset passwords without the need of human operators.
Emergency broadcast messaging	TTS+Speaker Verification	The voice can be biometric password to identify the broadcasters in combination with TTS technology that delivers notifications within the organizations.

Table 8 Employee Productivity Applications (Nuance Communications, 2010)

Transcription

The transcription function enables the creation of documentation from speech are widely used in medical and legal field. Traditionally, for health care field, medical transcriptions are typed by human medical transcriptionists. By implementing ASR technology, transcriptions can be created by computer and reviewed and edited by human medical transcriptionists. As well as in legal field, documents about the clients and cases can be created by ASR technology rather than requiring the lawyers to type by themselves. The application in transcription field is: (Nuance Communications, 2010)

Application	Technology Used	Details
Transcription / Documentation	ASR	Transcription or documentation can be created for medical transcription or legal documents. The computer-aided transcriptions can be reviewed and edited by human after the transcriptions are created. This is significantly productive since it is

		faster to edit than type.

Table 9 Transcription Applications (Nuance Communications, 2010)

3.3 Conclusions for Thailand's Opportunities

As for Thailand, the applications that will be suitable for Thailand must link to the level of capabilities of the speech technology in Thai language as well. For TTS technology, the capability is considered to be acceptable and is able to read almost all domains in Thai language. However, for ASR technology, the domain is limited to only sentences (for NECTEC is currently limited to only isolate words). Each area of the speech applications in previous section will be analyzed for the possibilities of implementing them in Thailand as follows:

Accessibility

Accessibility applications are one interesting area of speech applications due to their wide-ranging nature. Any applications that help people access to information and data in an easier way can be categorized in this area. The accessibility applications for blind people are already in use in Thailand such as screen reader. More research and development and knowledge transfer to the customers can be increased in this area. As for accessibility functions in office software, which can be used with normal people as well, it is an interesting application to go for due to the increase in competitions among the software developers. The accessibility functions can create value-added gimmicks for the office software. However, the dictation by speech is not yet ready for Thailand due to the limited capability of ASR technology, which is limited to only sentence level, not continuous speech.

Automation by Voice

Voice can be used as a remote control in situations that the users are not hands-free. This segment of the market can be an interesting market since people these days need to do things simultaneously, which require hands-free and eyes-free aided tools. Automation by voice can be interesting market for jobs that workers tend to be occupied with hands or eyes such as in hospitals. However, in some circumstances, for example in factory, automation by voice would not be suitable due to the noisy environment, which will distract the recognizer. Home appliance or media devices can also implement speech technologies to create value-added function.

Automotive

For automotive applications, speech applications can provide value-add function to the cars. The nature of the speech applications are considered suitable to the automotive applications because they are hands-free and eyes-free. In Thailand, voice control function is embedded in some new lines of automobile, but the language used is English. Although the costs of implementing speech technologies in other countries are lower, there is still cost to it. Most automotive lines that implement speech technologies tend to be more luxury or have technological advancement image such as the more expensive model of the All-New Ford Fiesta that is just launched in Thailand. For the car technology enthusiasts in Thailand, voice control function is one value-added function that they see as innovative. In the future, after the commercialization of the English voice control, the responds from the customers might lead to the initiation of using Thai language speech technologies.

Call Center

The call center is a very interesting market segment for speech technologies due to the nature of the business that requires speech as the main component. Also the growth rate of the call center business in Thailand is very attractive to invest in. According to Frost & Sullivan Thailand call center market is forecasted to hit a year-on-year (YoY) growth of

18.4% in 2016. (Frost & Sullivan, cited in The Nation Newspaper, 2010) The speech applications can be used to answer general inquiries, which will improve the productivity of the call center. The waiting time for the callers will decrease since there is no need to wait for the human operators. The human operators can be reserved for more complex tasks that cannot be managed by the speech applications. Call center is an interesting business in Thailand since there is still a lot of space for improvements.

Mobile Applications

Mobile applications are another field that speech can be applied to due to the nature that supports the speech applications. The mobile application developers or service providers can implement speech applications such as mobile functions or mobile downloading services. The mobile market in Thailand has interesting opportunities to invest in. According to Thailand ICT Market 2009 and Outlook 2010, the wireless equipment market in Thailand in 2009 made up the market worth of 23,925 million THB or approximately US$ 750 million with the annual growth of 4.3%. Also, the mobile service market in 2009 made up the market value of 177,589 million THB or approximately US$ 5,550 million. (NECTEC, 2009) The investment in mobile business is interesting choice due to the large market expansion in Thailand in the previous years, and speech applications can be value-added features that provide value to the customers.

Employee Productivity

The employee productivity suit can be used in companies to increase a company's workforce productivity. The employee productivity suit can be used to ensure consistent in business processes and employee communications, and therefore decrease the IT costs. The speech applications might be suitable for large companies to invest due to their high IT

costs. Automated attendant and password reset might be suitable for companies in Thailand.

Transcription

The transcription applications are not ready to be implemented in Thailand due to the constraint in the capabilities of ASR function. The level of domain that the transcription applications require is continuous speech, which is currently not capable with ASR technology of NECTEC or other existing providers in Thailand.

Chapter IV
Market Study of Current Potential Market

4.1 Status of Speech Technologies in Thailand and its Potential Areas

The Thai language speech technologies have made the progress in the research and development slowly comparing to other language such as English due to the complexity of its nature. According to Wutiwiwatchai and Furui, Thai language is a tonal language with no explicit word boundaries, which makes the research and development on Thai speech technologies challenging. (Wutiwiwatchai and Furui, 2006)

According to NECTEC and the preliminary study of the speech technologies business environment in Thailand, speech technologies for Thai language have been developed for a long time at NECTEC but have not been widely commercialized due to the constraints in technical aspects of the technologies. However, currently the TTS technology software, VAJA 6.0, has been confirmed by NECTEC of the efficiency and effectiveness that it is ready to be commercialized. As for the ASR technology software, iSpeech, its readiness for commercialization is expected to be in the next year or two. For ASR technology that is planned to commercialize, the level of its capability is limited to limit domain, which can input only isolated words. The status of speech technologies in Thailand is considered an emerging technology, very new, and is considered to be an innovation.

The potential areas in the near future of speech technologies in Thailand that have demand from the software developers according to the preliminary study are call center, accessibility, automation by voice, and mobile phone applications, and e-book reader. Since there are demands of speech technologies from these areas and they are considered successful in other countries. However, the business environments of other countries are different to that of Thailand; therefore, the study of Thailand's business environment and NECTEC's capability assessment must be conducted. The speech technologies can either value add and create innovations, which lead to the increase in customer satisfaction, or

increase efficiency and effectiveness of the working process, which lead to the decrease in cost.

4.2 Speech Technologies Providers in Thailand

4.2.1 Overview

The speech technologies are considered new technology for Thailand, and high investment cost must be covered for the research. In Thailand, only several sectors are involved in developing and commercializing Thai language speech technologies in basic technology form. According to Frost & Sullivan, Nuance Communications, the world leading speech technology provider, is currently working with local vendors in developing Thai language technology. (Frost & Sullivan, 2008) The following analysis of the current speech technologies providers in Thailand is categorized in terms of technological advancement and revenue from commercialization. The information on this analysis is based on an interview with NECTEC's researchers that have been in speech technology business for a long period of time.

*** It is noted that the number is the approximation from the experience of NECTEC's researchers not the exact numbers. The number accounts only for Thai language speech technologies (TTS and ASR) in Thailand.

4.2.2 Text-to-Speech (TTS) Providers

There are currently three main TTS technology providers in Thailand: NECTEC, Chulalongkorn University, and private companies such as PPA Innovation Co.,Ltd. and Sun Systems Corporation Limited (collaborate with Nuance Communications).

TTS Technological Advancement

In terms of technological advancement, NECTEC claims to be the leading sector in developing TTS technology due to its long period of research and development and the support from the public sectors. The latest version of TTS software by NECTEC, VAJA 6.0, has been accepted as the advanced TTS software at the moment. However, due to the nature of the organization that is different from those of private companies, NECTEC has been devoted in intensive research and development area but has not been involving much in the commercialization area.

Chulalongkorn University is another sector that is similar to NECTEC in the way that much of the efforts are in the research and development area but not much commercialization effort has been made. Many sources of funding from the society are provided to Chulalongkorn University to increase the research knowledge that will eventually benefit the society as a whole. Since Chulalongkorn University is the public university that is devoted to the Thai society, the objectives of the researches are mainly to benefit the society.

The technological advancement of private companies is growing since the large international providers are joining the local vendors in research and development. PPA Innovation Co., Ltd. is a Thai private company operating in the artificial intelligence business field, and one of their registered products is the TTS software. PPA Innovation is one of the first private TTS technology providers in Thailand. Sun Systems Corporation Limited is a serviced solution provider, which collaborates with Nuance Communication in developing Thai language speech technologies to embed in their call center solutions.

TTS Revenue from Commercialization

According to Figure 14, when looking at the revenue from commercialization aspect, NECTEC does not gain much in terms of revenue. As a national research center, NECTEC does not focus on aggressive marketing strategy, but rather focus on partly

commercialization and partly supports to the society. The revenue share from TTS technology of NECTEC is approximately 30% of the total.

In the same way, Chulalongkorn University mostly put effort on their research and development program with the objective of benefiting the society. Their TTS software has not been commercialized extensively. Their revenue share from TTS technology is approximately 10%.

The situation is opposite for private companies in the revenue aspect. Their share of revenue took over 60% of the total revenue in TTS market. One of the largest customer of PPA Innovation is the “Thailand Association of the Blind”, which had bought the license from PPA Innovation for their use.

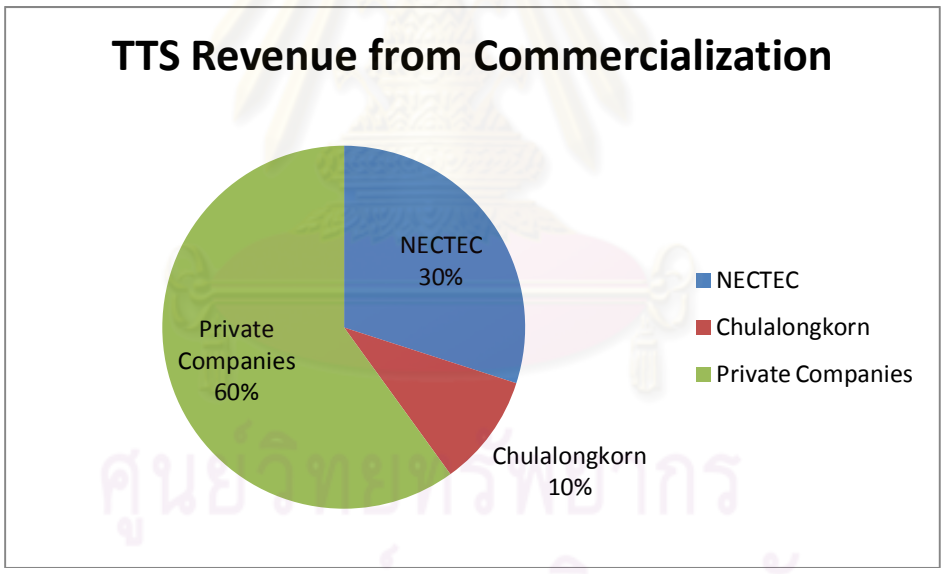


Figure 14 TTS Revenue from Commercialization (NECTEC, 2010)

**Note: The number is only the approximation from experience.*

4.2.3 Automatic Speech Recognition (ASR) Providers

There are currently three main ASR technology providers in Thailand: NECTEC, Chulalongkorn University, and private companies such as Tell Voice Co.,Ltd. and Sun Systems Corporation Limited (collaborate with Nuance Communications).

ASR Technological Advancement

For ASR technology, the private companies are the leader both in technological advancement aspect and revenue from commercialization. Tell Voice is a Thai company that has been in the research and development field of ASR technology for over a decade. The company developed the basic technology of ASR and provides ASR technology to their customers in the form of ASR call center solutions. Similarly to Sun Systems, the company developed ASR technology in collaboration with Nuance Communication to support their call center solutions.

NECTEC's ASR technology is less advanced than its TTS technology, and is not yet ready to be commercialized at the present. However, the technology should be ready to be commercialized in the coming year or two.

Chulalongkorn University will also be increasing their research and development in ASR technology in the future.

ASR Revenue from Commercialization

According to Figure 15, corresponding to the high technological advancement, the private companies also gained the highest revenue from the commercialization of approximately 70% share of revenue in the market. The private companies provide their customers with complete solutions rather than just basic technology.

NECTEC has gained about 20% of the total revenue. There has not been a serious commercialization of ASR yet, and also the technology is not ready to be seriously

commercialized at the present. However, if NECTEC increases the pace of their readiness for commercialization, they might gain more revenue share in the future.

Likewise for Chulalongkorn University, the ASR technology has not much been commercialized at the present. Their presence in the market in terms of revenues is approximately 10%.

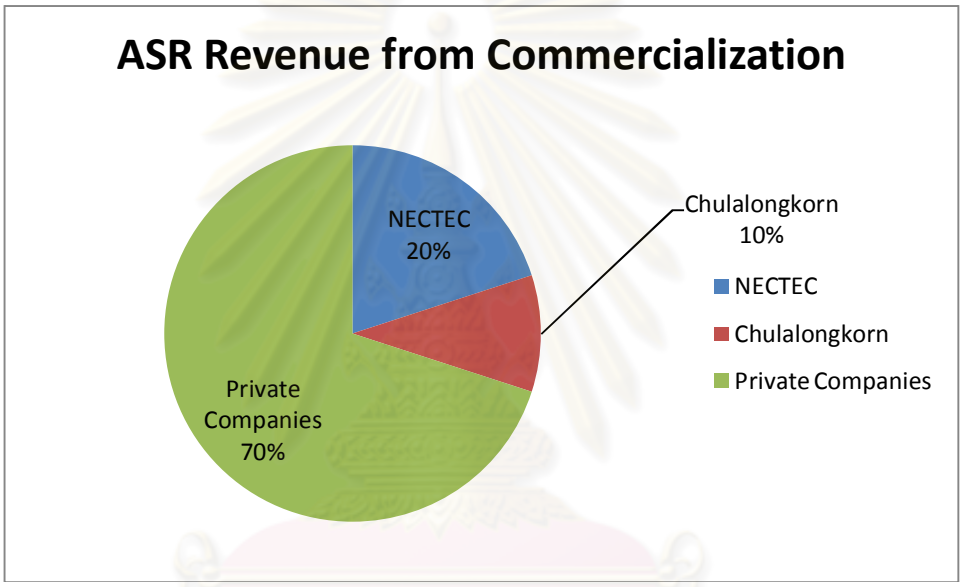


Figure 15 ASR Revenue from Commercialization (NECTEC, 2010)

**Note: The number is only the approximation from experience.*

4.3 Current Technical Capabilities of NECTEC's Speech Technologies

This part of the study is based on information from NECTEC.

4.3.1 Current Technical Capabilities of VAJA 6.0 (NECTEC, 2010)

VAJA 6.0 is Thai language speech synthesis software developed by NECTEC. The technology enables the conversion of Thai text into synthetic speech that people can understand. The installation and implementation is considered simple. The specific basic requirements to install VAJA 6.0 are as follows:

Supported Platform	• Windows XP, Windows Vista, Windows 7 • Linux
Hard Disk Size	Approximately 170 MB
CPU	1.6 GHz
Required RAM Size	Approximately 512 MB
Other Requirements	Speaker

Table 10 Basic Requirements for VAJA 6.0 (NECTEC, 2010)

The capabilities level of VAJA 6.0 can be determined in terms of speech quality, naturalness, and speech intelligibility.

- Speech Quality – The quality of the sound produced.
- Naturalness – The ability of the system to imitate the human's speech.
- Speech Intelligibility – The ability of the system to correctly produce the correct pronunciations that match with the words.

VAJA 6.0 delivers high quality speech that is similar to human's speech with almost 100% accuracy. VAJA 6.0 has the strength that exceeds other software that it can create almost any word in Thai language because the software has the text analysis function that can analyze even the words that are not found in the dictionary. Moreover, the users can even input the new specific words such as the names and authorize their pronunciations as well. The software can be used via the serviced websites or in package software form. The capabilities of VAJA 6.0 can be summarized as follows:

Speech Quality	High quality
Naturalness	Similar to human's speech
Speech Intelligibility	Approximately 97% accuracy
Domain	Any word in Thai language including words that are not

	found in dictionary.
Other Capabilities	Input specific words such as name to authorize the pronunciation.

Table 11 Capabilities of VAJA 6.0 (NECTEC, 2010)

According to NECTEC, each aspect of the capabilities is evaluated by an evaluation from users by allowing them to listen and give a score from 1-5. The evaluation was evaluated in terms of "mean opinion score" (MOS), which calculated the average of all scores. The evaluation was carried out in terms of qualitative characteristics: speech quality & naturalness, speech intelligibility, and the benefits from the software. The environment in which the evaluation took place was as follows:

- The users that took the evaluations were Thai native speakers.
- The users that took the evaluations did not take part in the development of VAJA 6.0.
- The evaluation was carried out with the real system.
- The criterion was ranged from 1 (very bad) to 5 (very good).
- The evaluation was carried out with 98 users – 64 men and 34 women.
- The users that took the evaluations were in a wide range of occupations.
- The users that took the evaluations were in a wide range of ages.

The result from the evaluation is as follows:

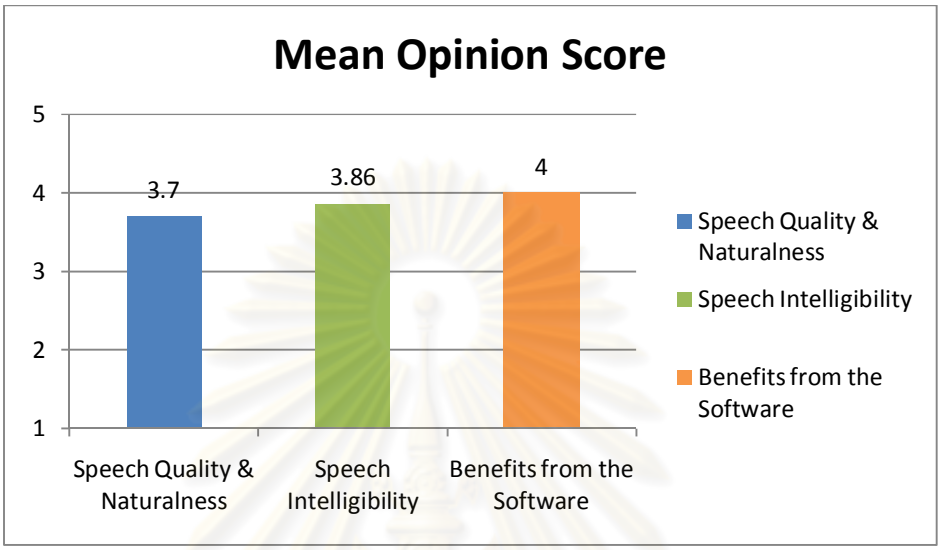


Figure 16 Mean Opinion Score (NECTEC, 2010)

From the result of the evaluation, it can be concluded that the users were considerably satisfied with VAJA 6.0 system. The capabilities in terms of speech quality, naturalness, and speech intelligibility are considered above average. Also, from the score of the benefits from the software, it can be concluded that the users perceive the TTS software as one useful technology.

4.3.2 Current Technical Capabilities of iSpeech (NECTEC, 2010)

iSpeech software is speech recognition system that can recognize the Thai language speech. The system converts spoken words into parameters that it can understand. The iSpeech system can be categorized by the level of domain that it can recognize. There are three main types of iSpeech:

- iSpeech- W: Isolated word recognition (IWR)
- iSpeech- R: Continuous speech recognition (CSR) with limited grammar
- iSpeech- N: Continuous speech recognition (CSR) with unlimited grammar

Currently, only iSpeech- W is ready in terms of technical capabilities to be commercialized. iSpeech- R and iSpeech- N is the is still in the research and development process.

iSpeech- W

iSpeech- W is the ASR system that can recognize isolated words such as street names, department names, and anything that are isolated words. The mechanism for iSpeech- W is that the words will be compared with the referenced words stored in the database and select the best-matched words. The users can edit and create the set of words that will match with their preferences. The specific basic requirements to install iSpeech - W are as follows:

Supported Platform	• Windows 2000, Windows NT, Windows XP • Linux
CPU	Pentium III or above
Required RAM Size	128 MB or greater
Other Requirements	Sound card that is supported by Windows OS

Table 12 Basic Requirements for iSpeech- W (NECTEC, 2010)

iSpeech- W is able to recognize the isolated words with more than 90% accuracy when used to recognize 100 words in an office environment. There are three main levels of domain as suggested above:

- Isolated Words- For example, voice command.
- Continuous Speech with Limited Grammar- Usually uses in spoken dialog system that can recognize a phrase or sentence that is within the selected scope.

- Continuous Speech with Unlimited Grammar- Any possible conversation likes human's conversation such as dictation and transcription.

For the working environment, there are three main levels of environment:

- Quiet Room
- Office
- Outdoor

Currently, iSpeech- W can be used within the quiet room to office environment. The capabilities of iSpeech- W can be summarized as follows:

Accuracy	More than 90% accuracy when used to recognize 100 words in an office environment.
Domain	Isolated words
Working Environment	Quiet room to office environment
Number of Words Able to be Recognized	Can input the unlimited number of words, but the accuracy is also decreased proportionally.

Table 13 Capabilities of iSpeech- W (NECTEC, 2010)

4.4 SWOT Analysis of NECTEC's Speech Technologies in Thailand

The SWOT analysis aims to identify the strengths, weaknesses, opportunities, and threats of NECTEC's Thai language speech technology (TTS and ASR) software in Thailand. The information to analyze SWOT is from NECTEC combining with the information from the preliminary interviews of NECTEC's customers and some other information from journals and news.

Strengths

- Easy to use
    - Speech is the easiest means of communication for humans; therefore, the technology will enable the use of technology to a broader range of people for example handicapped or people without computer literacy. Speech technologies will enable some functions that require input by keyboard or telephone keypad to be commanded by speech, which is easier for humans to use.
- Hands-free and eyes-free
    - Both TTS and ASR technology require no use of hands or eyes, which can be advantageous in many ways. The tasks using these technologies can be performed at the same time with other tasks, for example when driving.
- Fast and convenient
    - Speech is the speediest means of communication. By using ASR technology, speaking is about 3-4 times faster than typing on keyboard, also it is more convenient than input on telephone key pad. As for TTS technology, listening is easier than reading as well.
- High technical reliability for TTS
    - TTS technology by NECTEC (VAJA 6.0) has the strength that it can create almost any word in Thai language because the software has the text analysis function that can analyze even the words that are not found in the dictionary.
- Supports popular platforms
    - The software developed by NECTEC supports Windows and Linux platform, which are the most popular platform used in Thailand, and it will support

mobile platform by next year as well. Therefore, most of the software houses that are interested of speech technologies would be able to implement the technologies.

- Ease of implementation of TTS
    - The implementation of TTS is considered to be easy to implement.
- Speech technologies is considered value-add function
    - Speech technologies can be used to add value to existing products and services because the technologies are considered new and innovative for Thailand. Speech technologies can be used to create gimmick or add-ons for products or services.

Weaknesses

- Software's memory requirement not suitable for some platforms
    - The software processing takes a considerably large size of memory requirement when used with some platforms such as mobile platform. NECTEC is now developing the speech software that will be suitable with mobile platform, which will be ready to be commercialized approximately by next year.
- Accuracy issue of ASR
    - Currently, the speech recognition technology in Thai language is constrained to limit domain, which limit to only input of isolated words. For isolated words, the accuracy is considerably acceptable (more than 80% accuracy), but for longer speech input, the accuracy is not in the acceptable

range yet. Some customers still doubt the accuracy issue and are not confident enough to use ASR.

- Complexity of implementation of ASR
    - Some software houses believe that the implementation procedures are complex for ASR technology. However, the implementation is not as complex as they thought; therefore, knowledge transfer must be conducted to increase correct understandings between the technology providers and the customers.
- Customers are not confident to invest on speech technologies
    - Customers see that by implementing speech technologies, investments in many aspects other than licensing are required such as hardware, system, and labor. Therefore, as it is very new for Thailand, the customers are not quite confident to invest in it.
- Speech technologies are considered value-add function
    - As it is strength, it is also a weakness as well. Speech technologies are considered value-added function rather than core function. Most of the applications that use speech technologies use them as value-added functions or gimmick more than as core function. Therefore, the customers of NECTEC (software houses) might not see the importance or urgency to acquire speech technologies now. They might wait until the technologies become more abundant or even become a market requirement to start implementing it.

Opportunities

- International trend
    - The international trend from other countries such as the United States, the United Kingdom, or Japan can impact the direction of the technologies in Thailand. More applications in form of hardware and software from other countries embed speech technologies, for example, voice dialing function in Blackberry, voice command in Windows, etc. Therefore, Thai people will start to slowly get used to speech technologies and will demand more of them in the future.
- High global market growth rate
    - The annual global market growth rate of speech technologies are about 30% (Eastwood, 2005). Also, Asia Pacific is considered an emerging market for speech technologies.
- Customers pay more attention to innovation
    - Nowadays, customers in Thailand pay more attention to innovations, which can be seen from the very high demand of innovative products such as Iphone. If the customers pay more attention to innovation, it would persuade the application developers (software houses) to embed more innovative technologies. Speech technologies could be an interesting choice for innovation.
- Economic recovery
    - Thailand's economy started to recover at the fourth quarter of 2009 according to the Thai Industry Sentiment Index of 113.6 on December 2009, which is the highest in five year period (Thailand FTI, 2010). The recovery of

Thailand's economy leads to the new investments in order to replace old software and to create growth for the businesses (most businesses in Thailand stopped or slowed down investments since 2008 due to the economic recession). Businesses might be interested to create new business ventures that require new and innovative technology, and therefore, speech technologies can be one interesting choice. Even though since the coup on March and flooding situation in Thailand had made 2010-index dropped, overall the index is still in the recovery period comparing to 2007-2008.

Figure 17 Thai Industry Sentiment Index (Latest on October 2010)

(Thailand FTI: www.fti.or.th, 2010)

Threats

- Customers' unfamiliarity with speech technologies
    - The users in Thailand might not be familiar with speech technologies since they are very new. The unfamiliarity might create frustration for users when they could not get the result that they wanted. In the starting period, the combination of keypads and voice might be used to increase the familiarity.
- Political instability
    - The political situation in Thailand has not been stable for the past several years. Companies might not be confident in investing in new technologies since the investment-related policies can be changed if there is dissolution of the parliament.

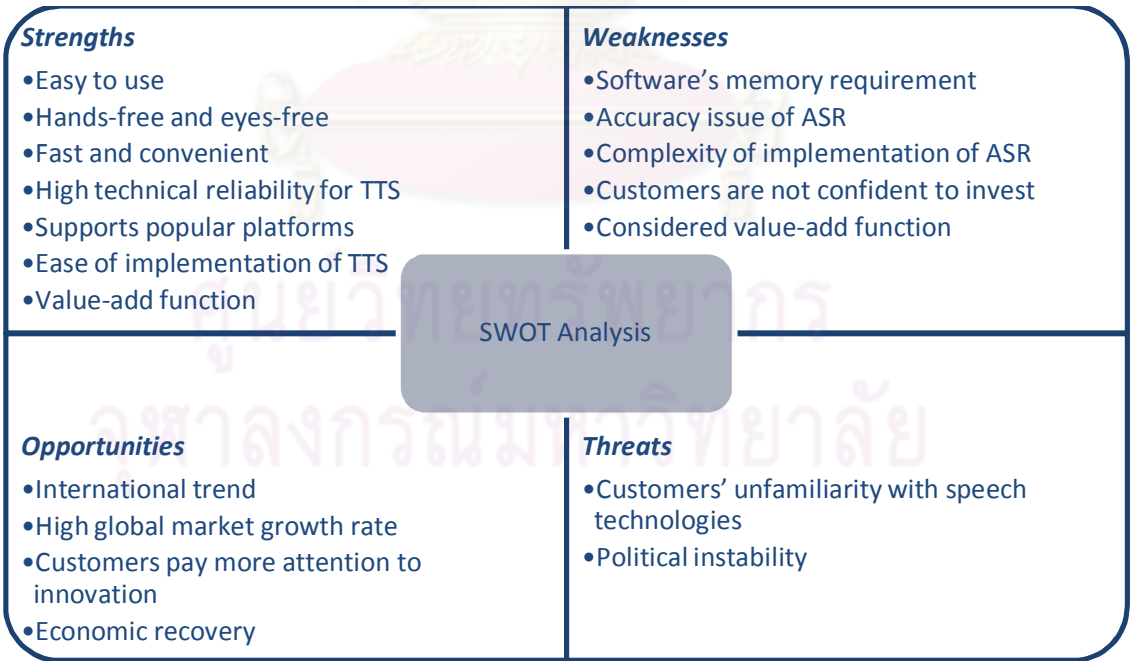


Figure 18 SWOT Analysis of Speech Technologies

4.5 Porter's Five Forces Analysis

This part of the study aims to analyze the attractiveness and the profitability of the speech technology's industry in Thailand by implementing Porter's Five Forces Analysis. The five forces suggested by Porter are: customers, suppliers, industry rivals, potential entrants, and substitute products. If these forces are strong, the industry's attractiveness and profitability are low. However, if these forces are weak, there are more chances of the earning profits in the industry. Therefore, it is crucial for businesses to find their position in the industry that can overcome these forces. The industry scope is only limited to the basic technology of speech technology (JAVA 6.0 and iSpeech software by NECTEC).

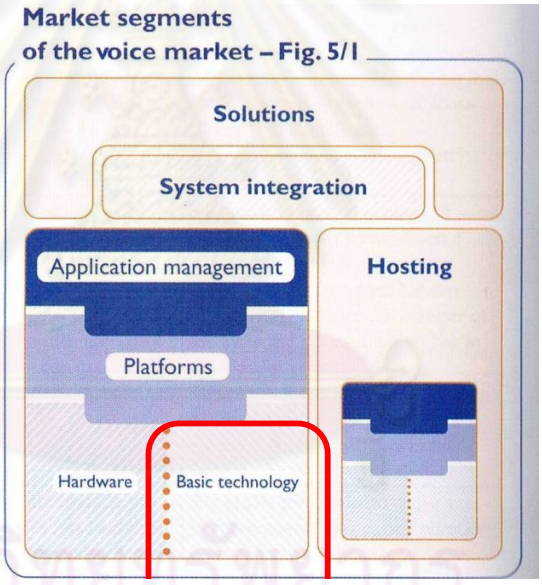


Figure 19 Value Chain of Speech Technologies (Artelt, 2008)

From "Voice Compass: International 2008/2009: Speech Goes Mainstream"

The five forces are analyzed as follows:

Threats of Entry

There is considerably low number of potential competitors in the industry as suggested in the previous section. The threats of entry can be analyzed in two aspects: the barriers to entry and the reactions of the existing players in the industry.

The barrier to entry in the basic speech technology industry is high due to numbers of reasons:

- To develop the basic technology of speech technology, the high investment both in terms of capital and research is required. The cost of research for speech technology is very high due to the complex nature of the technology. The intensive and long-period research is required; the research at NECTEC has been carried out for over ten years. Moreover, in terms of the human resource capability required, the training for specialized programmers for this field must be carried out. If the companies are not large enough in both funding and research background, the possibility of entering this industry is low.

- The economy of scale for the research is required in developing basic speech technology. If the research is not excessively carried out, the technology will not be acceptable in terms of quality. Also in terms of high investment cost, the economy of scale for the research is required.

- The customer switching costs for the basic technology is moderate. If a developer decides to choose one company's basic technology to develop a solution, there are some effects if the developer decides to change the vendor. The costs can occur from the related system integration that must be adjusted along with the change. However, the switching of the vendor is possible.

- There can be advantage in terms of brand and image of the technology providers. However, it is only significant in terms of local or multi-national vendors. Some

companies will prefer to use software from multi-national vendors due to the strong image of quality.

Other than the barriers to entry, the assessment in terms of how the existing players in the industry will react must be carried out:

- The industry is very new in Thailand and there is no clear competition in the industry. Since there are only several providers in Thailand, the industry have not yet form the direct competition against each other because the market is still in the developing stage.

Since the barrier of entry is high for basic speech technology industry, **the threat of entry in this industry is low**. This means that it is hard for new competitor to enter this industry mainly due to the high level of required initial investment cost and research. The company must be large enough both in size and power to invest in the high cost research and development program.

The Power of Suppliers

Powerful suppliers can lead to the limit profit potential because of the high costs that are not covered. Since there is no supplier for the development of speech technology software, the supplier in this context refers to the labour required for the development, which are the programmers. **The power of supplier in this case is considered moderate to high.**

- Due to the required skills and training in order to train the programmers for this specific field in speech industry, there are costs and time that must be invested on. For NECTEC, there are educational funding scheme for programmers that work on speech technology development to attract the skilled programmers. However, for

private companies that do not have enough funding or benefits to attract the programmers, the programmers might be able to negotiate with the companies.

The Power of Buyers

The power of buyers can be assessed in two main aspects: bargaining leverage and price sensitivity. The current bargaining leverage of the buyers is moderate due to several reasons:

- Since the technology is still very new, the customers (software developers) also take risks of implementing the technology since there is not much example of the applications of speech technology in Thailand. Therefore, the software developers can negotiate since they must also take risks of commercializing new technology.
- The potential customer groups in the first period of commercialization are still very limited since not every type of applications / solutions are suitable with using voice and not much technology transfer has been done in Thailand.
- However, the basic technology software is differentiated from each other by the quality of the technology such as speech quality, naturalness, speech intelligibility and accuracy. Since the products are differentiated, the customers tend to choose the better quality software.
- Also, it is quite hard for the software houses to backward integrate and develop the basic technology by themselves due to the high investment in capital and research. Or even to collaborate with international vendors in development, the investment cost is considered high.

However, the basic speech technology software is not price sensitive.

- The customers will give high importance to the quality of the technology since they must use the technology with their products or services; therefore, the quality of the basic technology will directly affects the quality of their products or services. However, it is noted that if there is a very high gap between different brands of software (for example some software provided by the local vendors sometimes has much lower price than the multi-national vendors'), the customers tend to go for the most cost-effective solutions.

In conclusion, the bargaining leverage of the buyers is moderate while the industry is not price sensitive. Therefore, **the power of buyers in this industry is considered moderate to low**.

The Threats of Substitutes

The threat of substitutes for speech technology is moderate to low, since speech technology is actually the substitute of the existing touchtone (keypads or buttons system) technology. Some customers might persist in using the existing touchtone system, but the future trend according to other countries is moving toward speech technology. For now, the only substitute of speech technology is touchtone system, which is already existed and might be outdated in the future.

Rivalry Among Existing Competitors

The rivalry among existing competitors directly affects the profit potential. If the rivalry is high, the profit potential tends to be low. The rivalry can be assessed in two main aspects: the intensity of the competition and the basis on which they compete. The intensity of the competition is low for basic speech technology industry.

- There are very few competitors in the industry, and the competition is not strong now since the commercialization of the technology has not been intensively initiated.
- Two of the three provider sectors in both TTS and ASR technology are not private companies (NECTEC and Chulalongkorn University) and both sectors are not involved in aggressive marketing. As for the private companies in the industry, the private companies are stronger in terms of marketing.
- The speech technology industry competes in terms of quality rather than price-cutting since the buyers of the industry are not price sensitive (in case the prices are not extremely different).

In conclusion, **the rivalry among the existing competitors is low**.

Conclusion on Porter's Five Forces Analysis

The five forces can be summarized as follows:

- Low threat of entry
- Moderate to high power of supplier
- Moderate to low power of buyer
- Moderate to low threat of substitutes
- Low rivalry among the existing competitors

In conclusion, the basic speech technology industry is considered attractive, but the sectors that want to enter this industry must be large enough in size, power, and capital investment to initiate adequate research and development. The companies that want to enter this industry must be aware of the point that speech technology is only one part of a solution; therefore, the implementation is limited to only some suitable solutions. The basic technology providers might also have to forward integrate and provide the solutions to the

end users to increase the customer base. Also it should be noted that the human resource for developing the speech software is limited and requires training, which the companies must have enough funding and strong research team to contribute in research and development. The competition is still not strong currently because the speech technology is very new in Thailand with only few sectors contributing to it.

4.6 Data and Analysis from Market Research

4.6.1 Detailed Market Research Procedure

The market research of this study has been carried out during October 2010 to February 2010 by interviewing the software houses, which are the potential customers of NECTEC. The software houses are the potential customers that might acquire license of the basic speech technologies and develop into applications / solutions. The market research can be separated into three main parts: the preliminary interviews of various types of software houses, the deep interviews with the call center software houses, and the interviews with the end users of the call center software houses. The research was conducted with the focus on the interest and plan of implementing speech technologies and other related information in order to analyze the outlook of the speech technologies in Thailand.

The market research procedures can be summarized as follows:

1. Preliminary interviews of various types of fifty software houses had focused on their interests on speech technologies. The interviews aimed to indicate the type of the software that have demand for speech technologies, and it also aimed to analyze the opinions of various types of software developers on implementing speech technologies with their software. From the preliminary interviews, the analysis on the demand and suitability of speech technologies were carried out to identify the most suitable market for speech technologies. The interviews were carried out by phone

interviews. The interviewed companies were chosen randomly from the contact companies of NECTEC. There were total of 289 software companies that NECTEC in contact. The sample size is calculated base on Taro Yamane's theory with the formula: (Taro Yamane, cited in Determining Sample Size by Israel)

$$n = \frac{N}{1 + N(e)^2}$$

Where n is the sample size, N is the population size and e is the precision level. In this case, N is 289 and e is 13%, from this the sample size is 49. Therefore, 50 companies were chosen randomly from the list.

2. After the preliminary interviews were carried out, the conclusion on the target customer group for speech technologies was determined. The target customer group that NECTEC should focus on is the call center business. Therefore, in-depth interviews on ten call center software developers were carried out. Some interviews were carried out by phone interviews, and some focus group interviews with the key players in the call center business were carried out. The interview focused on the implementation of the speech technologies in call center business.

3. Five end user companies of the call center software in different industry verticals were interviewed to determine the actual demand from the end users. However, due to the confidential issue, only a very small part of the end user group was examined. The end users are the companies that buy call center solutions from the software houses.

4. The information gathered from the interviews in combination with secondary information gathered were processed and analyzed.

4.6.2 Overview of the Preliminary Interview

The survey was done on total of fifty companies. The companies were chosen randomly from the contact companies of NECTEC, which are software developer companies. The interviews aimed to indicate the type of the software that have demand for speech technologies, and it also aimed to analyze the opinions of various types of software developers on implementing speech technologies with their software. From the preliminary interviews, the analysis on the demand and suitability of speech technologies were carried out to identify the most suitable market for speech technologies. The interviews were carried out by phone interviews.

4.6.3 Companies Profile

Type of Company

All companies are all developer companies, but may be operating with other type of services as well. Each company may be described as one of the seven following types of company or the combinations of the seven types:

1. Developer
2. Importer / Sale Representatives
3. Retailer
4. Services
5. System Integrator
6. Other types of company not mentioned

Other types of company not mentioned include:

- Provider
- Consultant

According to Figure 20, most of the companies, 61% percent of total, are developers. It is followed by companies that provide services of 16%, and companies that are importer or sales representatives of 14%. The system integrators made up 5% of the total, and the retailers made up 1%. Other types of companies made up the left over 3%.

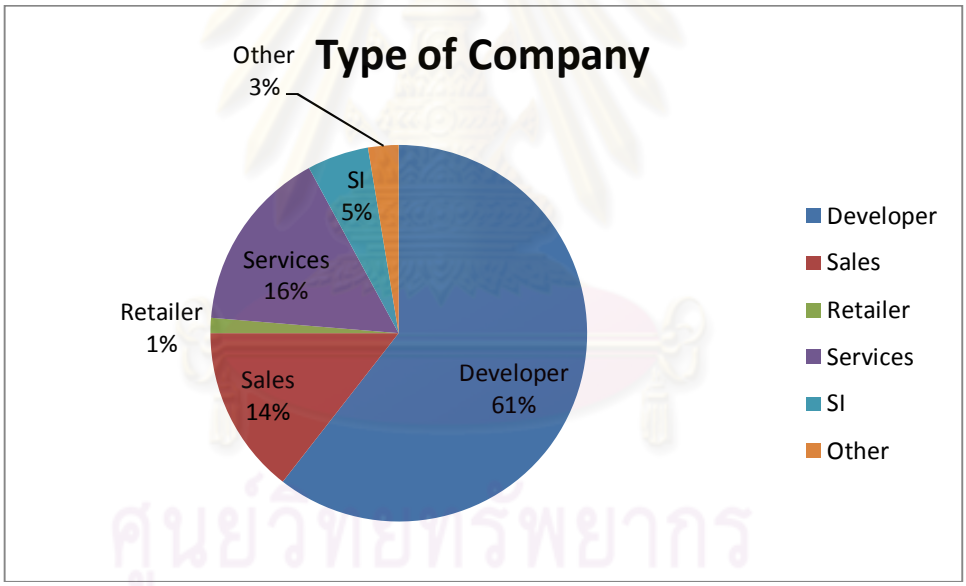


Figure 20 Type of Company

Type of Software

Each company might do business on one or more types of software in the following three categories:

1. Enterprise Software
2. Mobile Software

3. Embedded Software

According to Figure 21, enterprise software in this study is defined as application software that performs business functions, which is typically used in business, made up 63%. Mobile software is defined as application software that is used on mobile devices, made up 24%. Embedded software is defined as software for integral role with hardware (usually not for computer, but rather are for other machines), made up 13%.

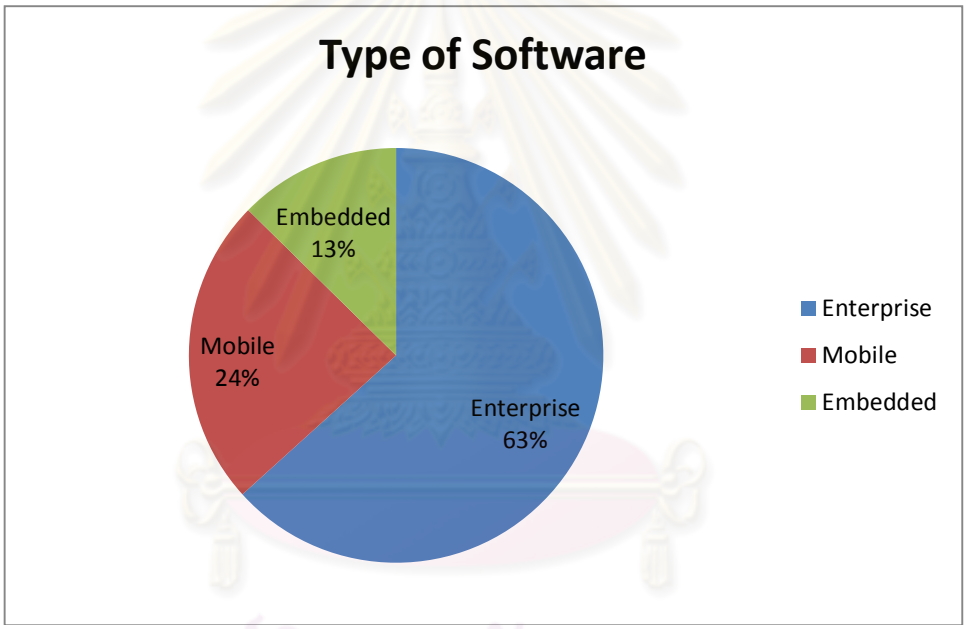


Figure 21 Type of Software

Business of the Company

Each company might do many businesses in different industries, where it can be categorized in one of the following categories:

1. Telecommunications (mainly call center)
2. Office & Administrative Software
3. Account Software

4. Manufacturing Software
5. Retailing Software
6. Information Management Software
7. Mobile & Portable Devices
8. Media & Game
9. Other industries not mentioned

Other industries not mentioned include:

- Banking Software
- Website
- Electronics Software
- Infrastructure
- Medical Software
- RFID system
- Online Map System

According to Figure 22, from the total number of companies chosen randomly for interview, the office & administrative software is the most abundant industry with 25% of the companies in this business. The second most abundant is the telecommunications business with 12% of the total. Whereas other businesses that are not in the list made up 17% of the total. Account software, mobile & portable devices, and information management system rank next on the list with 9%, manufacturing software with 7%, and finally media & game and retailing software with 6%.

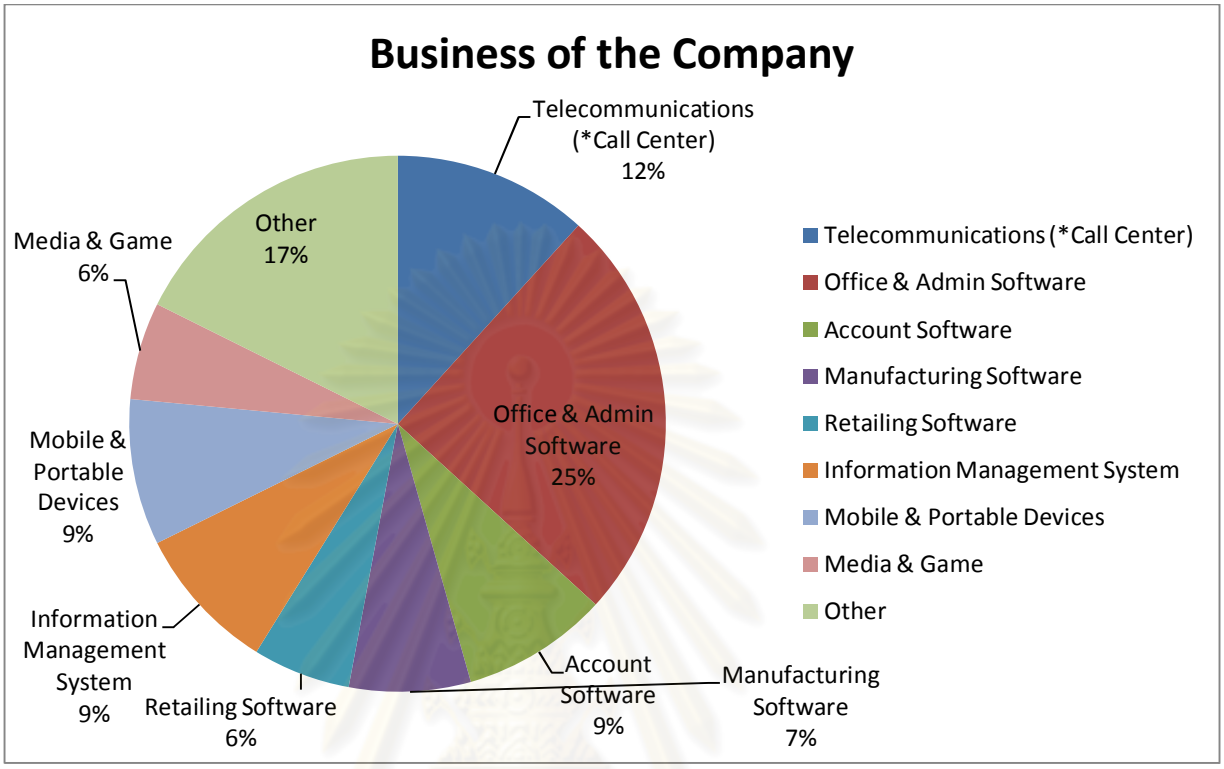


Figure 22 Business of the Company

Technology Platform

The technology platforms that the companies work on compose of:

1. UNIX
2. Mobile
3. Open Source (Mostly Linux)
4. RFID & Embedded Software
5. Windows

According to Figure 23, the main platform that the companies work on is Windows, which made up 59% of the total. The second most widely used platform is the open source, which is mostly Linux platform, made up 26%. Mobile, UNIX, and RFID & embedded software platform made up 9%, 3%, and 3% respectively.

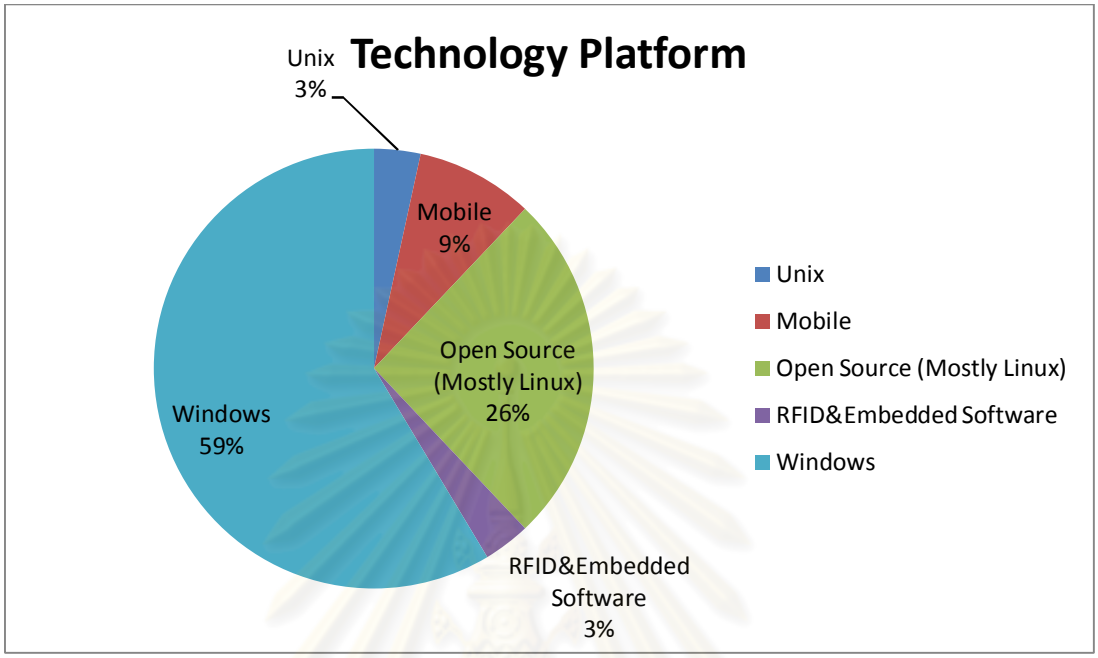


Figure 23 Technology Platform

4.6.4 About the Customer of the Companies

Customer Sectors

The customer sectors of the companies interviewed consist of:

1. Public Sector
2. Private Enterprise (medium and large size)
3. Small Enterprise
4. Consumer Market

According to Figure 24, the private enterprise which includes the medium and large size private enterprise made up 51% of the total customer base. The medium size private enterprise is defined by the number of computers used in the companies of 20 to 100 computers, whereas the large size private enterprise is defined by the number of computers used of over 100 computers. Public sector made up 27% of the total customer base, and small enterprise including family businesses made up 17% of the customer base. The

smallest customer base is the consumer market where the companies sell directly to the consumer, and it made up 5% of the total customer base.

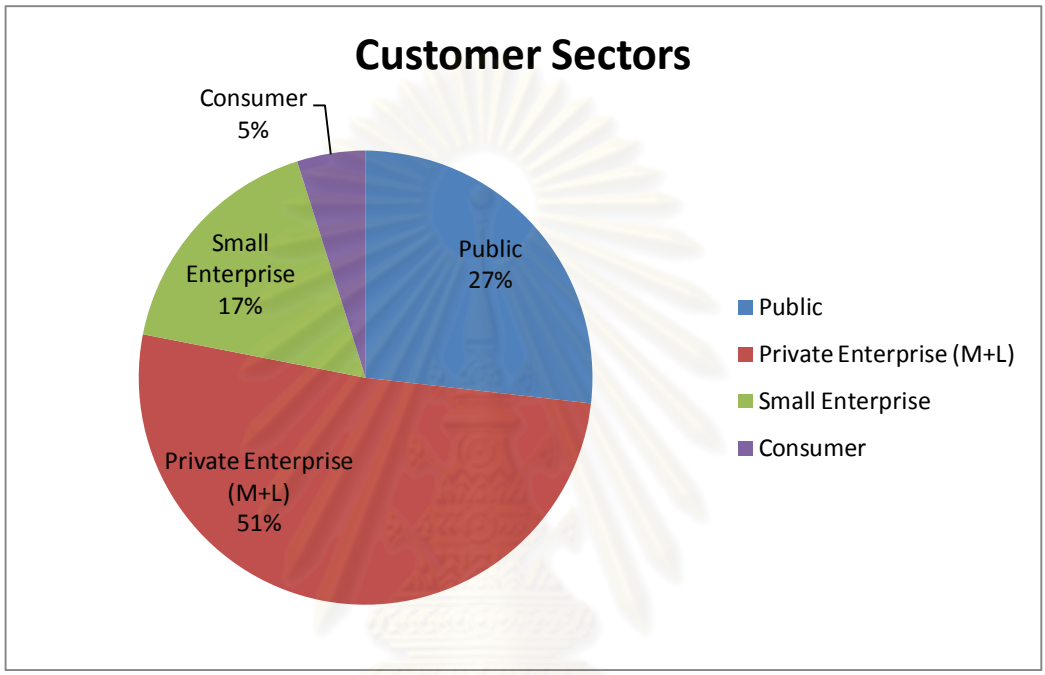


Figure 24 Customer Sectors

According to Figure 25, for the private enterprise, which compose of the medium and large size enterprise, the medium size enterprise made up 59% of total private enterprise, and large size enterprise made up 41% of total private enterprise.

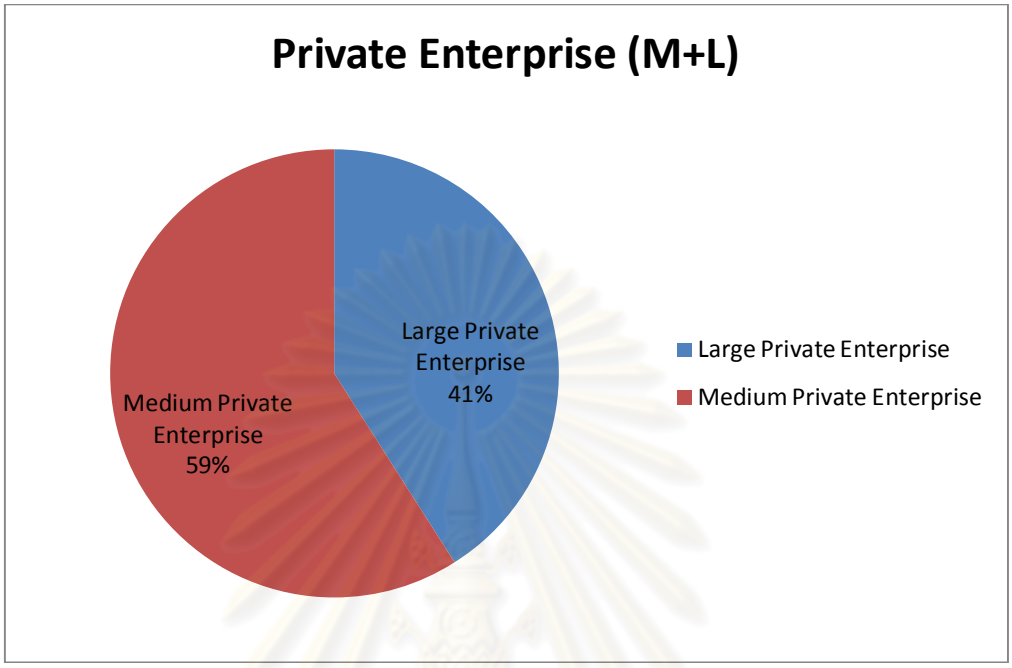


Figure 25 Private Enterprise (M+L)

Customer Industries

The industries of the customers of companies are in various sectors including:

1. Financial
2. Travel & Leisure
3. Logistics
4. Medical & Health Care
5. Education
6. Agriculture
7. Manufacturing (not including automotive manufacturing)
8. Automotive Manufacturing
9. Telecommunications
10. Retail

11. Other sectors not mentioned above.

Other sectors not mentioned include:

- Utilities
- Petroleum
- Technology
- Web-game
- Electronics
- Securities
- Real estate
- Insurance
- Wholesales
- Industrial retail
- Construction

According to Figure 26, manufacturing (not including automotive manufacturing), telecommunications, and other sectors not mentioned are each 12% of the total customer base. Financial, travel & leisure, and education sector are each 11% of the total customer base. Medical & health care, automotive manufacturing, logistics, and agriculture made up 7%, 6%, 4%, and 3% of the total customer base respectively.

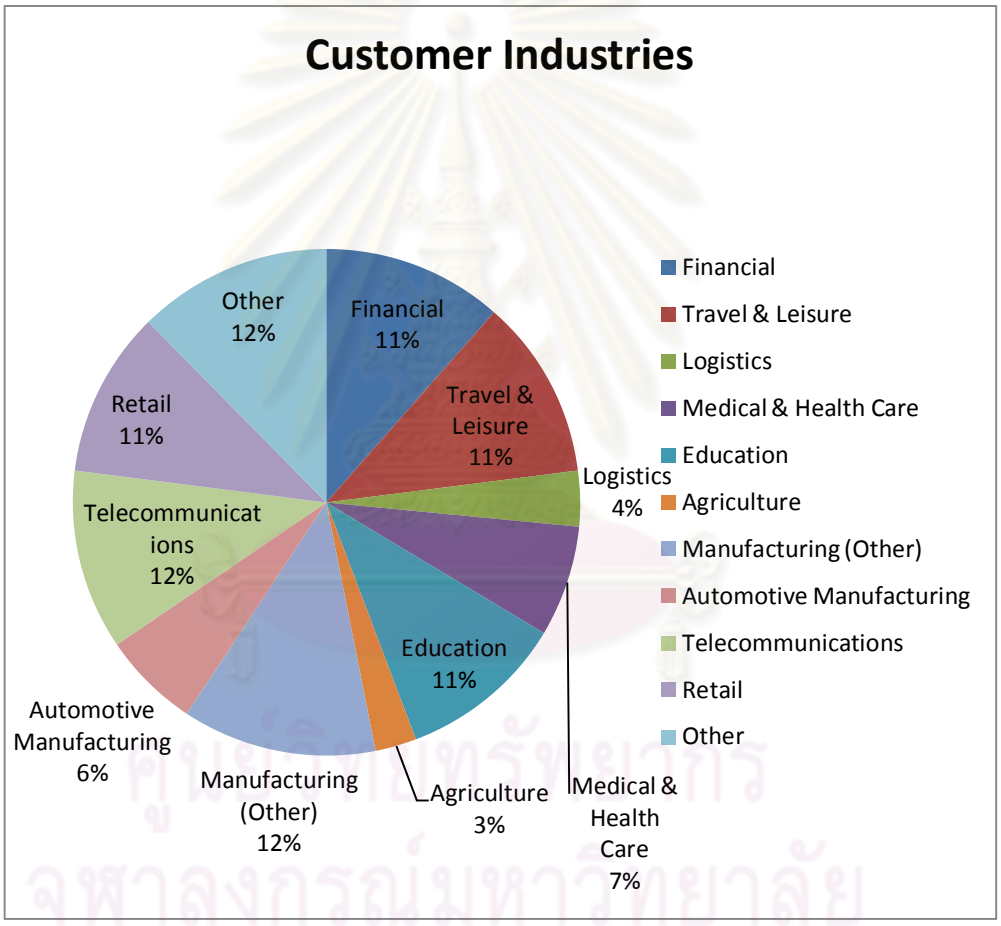


Figure 26 Customer Industries

4.6.5 Awareness & Interest in Speech Technology

According to Figure 27, the companies that are aware of and understand the use of speech technologies are 68% of the total companies surveyed. Whereas the companies that are not aware and do not know what speech technologies are 32% of the total companies surveyed. It can be seen that the percentage of unawareness is considered high because the interviews were conducted with the people who are in the technology field and they should know a lot about the technologies available.

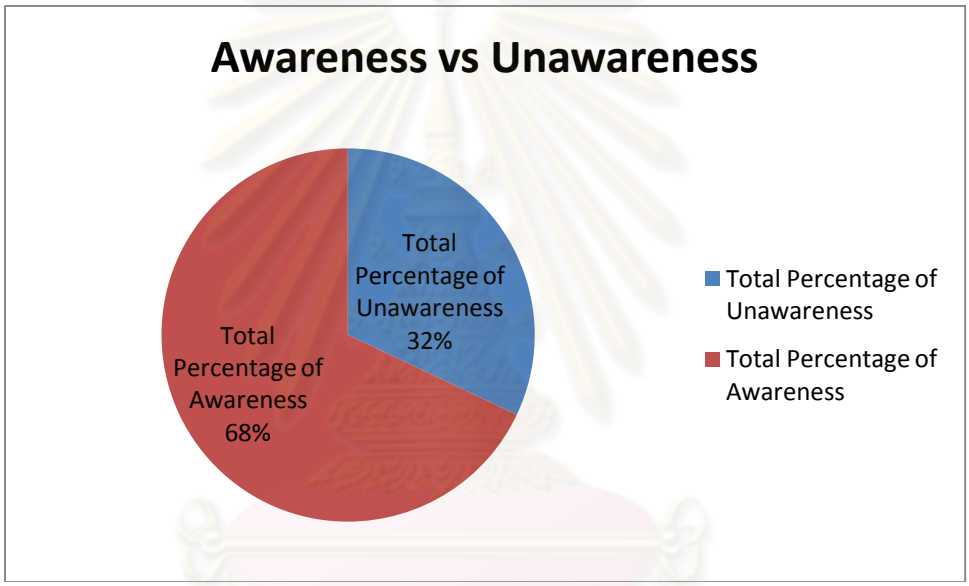


Figure 27 Awareness vs. Unawareness

According to Figure 28, from the total number of companies that are aware of speech technologies, only 41% of them have experience of using speech technologies. More than half (59%) never have experience with speech technologies. Most of the people who are aware of this technology have not even experience the technology; therefore, increase in road shows might increase the experience rate.

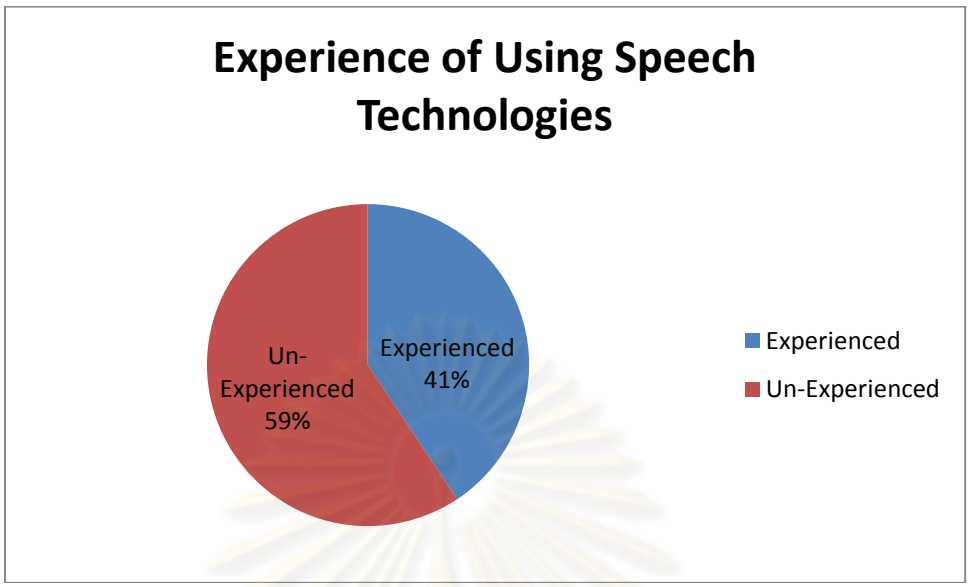


Figure 28 Experience of Using Speech Technologies

According to Figure 29, the interested companies made up 52% and the companies that are not interested made up 48% of the total companies surveyed. In this case, the rate of interest and not interest can be said to be approximately equal (about 50-50).

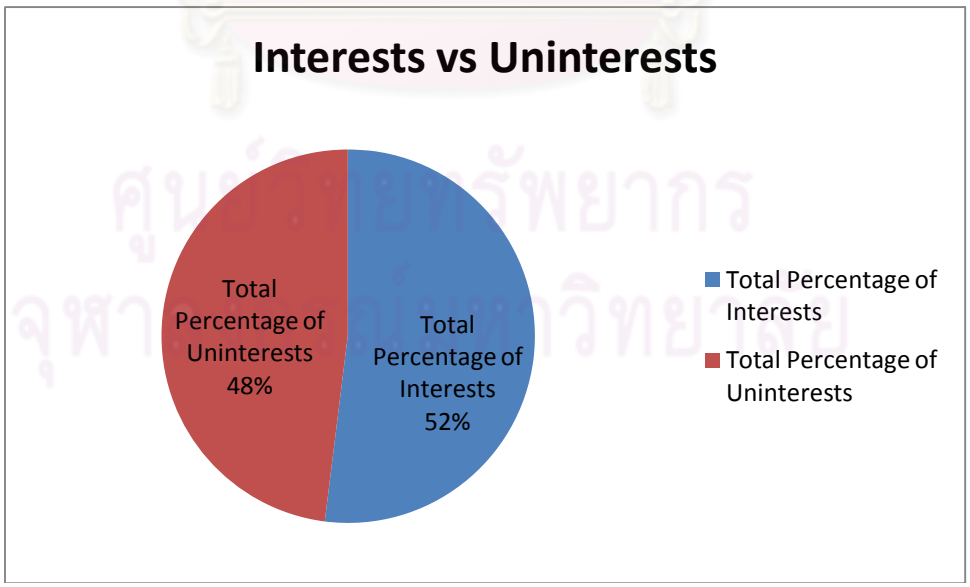


Figure 29 Interests vs Uninterests

According to Figure 30, among the companies that know and understand about speech technologies, 41% are very interested in speech technologies, 24% are moderately interested, and 35% are not interested. It can be concluded that more than half (65%) of the companies that have understandings about speech technologies will be interested. Therefore, the increase in awareness and experience of the technologies is important since understandings can lead to higher interest rate.

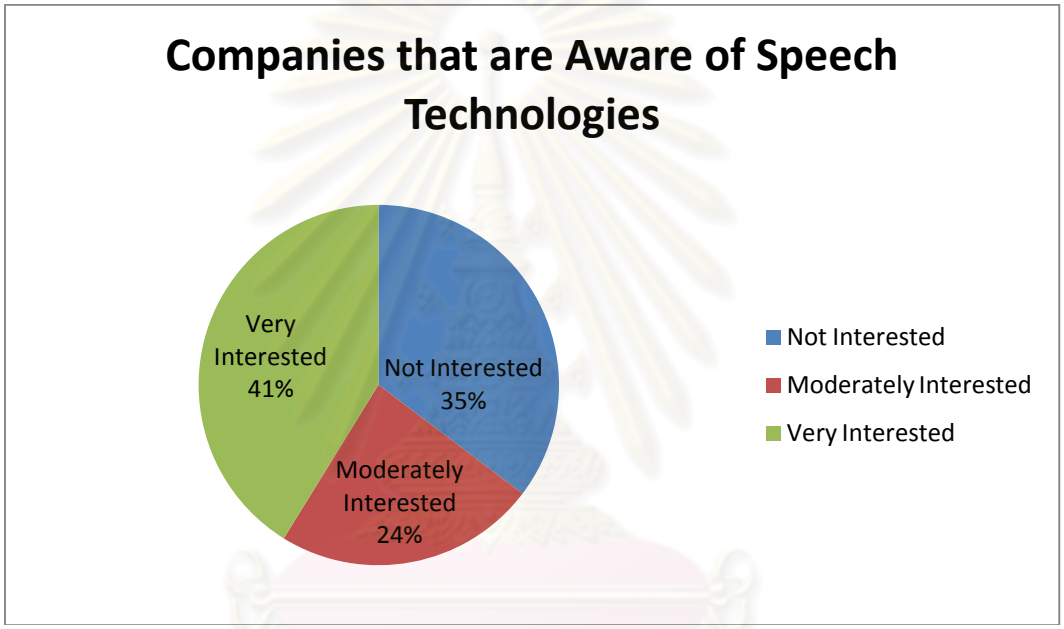


Figure 30 Companies that are Aware of Speech Technologies

According to Figure 31, for the companies that do not know about speech technologies, during the interviews the interviewer explained the basic concept of the speech technologies and ask if they are interested. The interested companies made up only 25% of them and 75% are not interested. It can be concluded that companies that know and understand about the speech technologies tend to be more interested. Therefore, the total interest rate might be increased by increasing marketing effort.

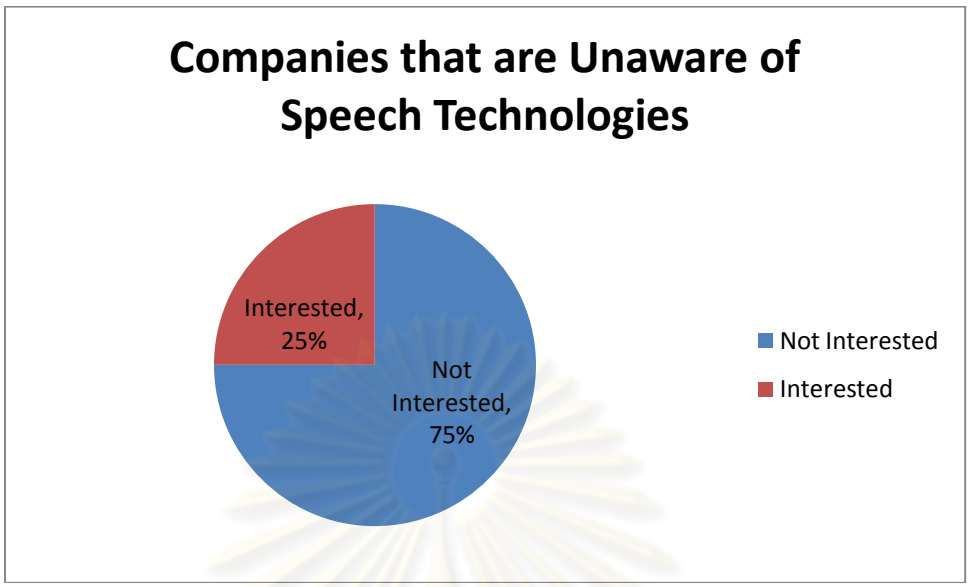


Figure 31 Companies that are Unaware of Speech Technologies

From the awareness and interest rate analyzed above, it can be concluded that there are five types of awareness and interest:

1. Aware of / Very Interested
2. Aware of / Moderately Interested
3. Aware of / Not Interested
4. Unaware of / Interested
5. Unaware of / Not Interested

According to Figure 32, the companies that are aware of the technologies and are also very interested made up 28%. The companies that are aware of the technologies and have moderate interest made up 16%. The companies that are aware of the technologies but are not interested made up 24%. As for the companies that are not aware of the technologies, the companies that are not aware of the technologies but have interest made up 8%, and the companies that are not aware of the technologies and are not interested made up 24%.

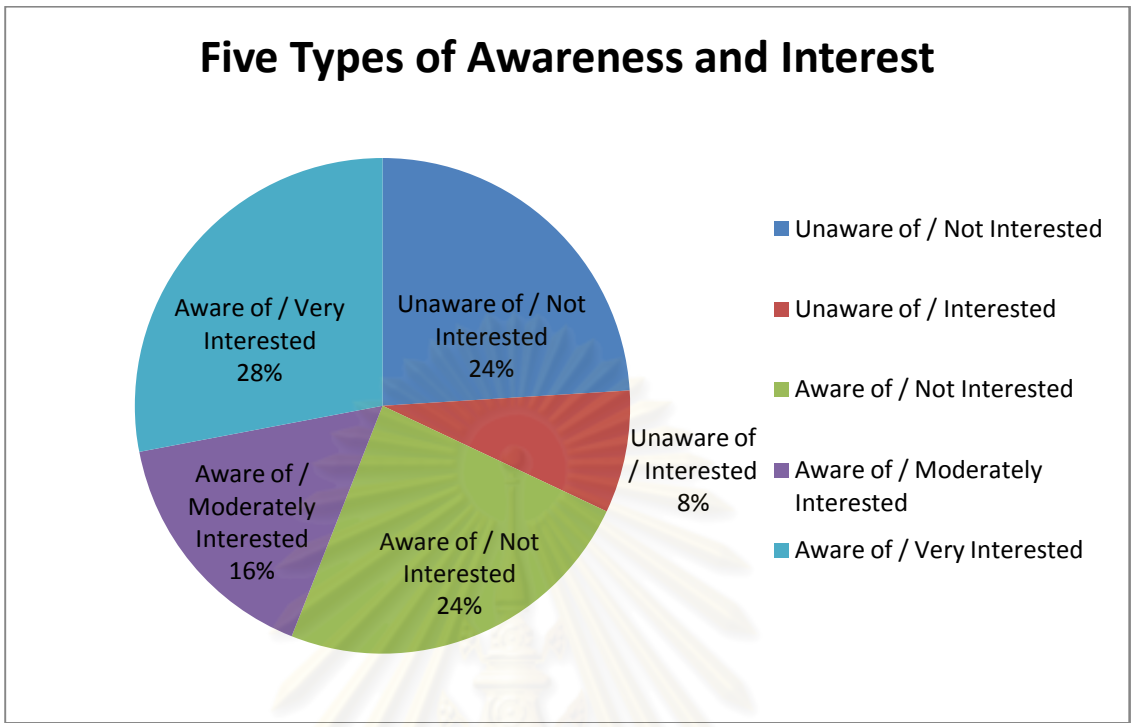


Figure 32 Five Types of Awareness and Interest

For each type, the solution to tackle each group is different. For the companies that are aware of the technologies and are also very interested, NECTEC should approach them first since they already have plans in their mind. This group is very important because they can be the technology pacesetters for speech technologies, and NECTEC should support them for the implementations.

For the companies that are aware of the technologies and have moderate interest, NECTEC should support them with application plans because these companies are already interested in using speech technologies, but they do not have the application plans yet. More direct contact to these companies in providing them the support for application plans can turn into real implementations

For the companies that are aware of the technologies but are not interested, it might be hard for NECTEC to approach these companies since they already know about the technologies but are not interested. However, if NECTEC really do not want to lose this customer group, vigorous efforts to create application plans for companies in this group is necessary.

For the companies that are not aware of the technologies, increase in marketing effort is needed. More effort can be placed at the interested companies since they can be more open-minded, and also application plans should be provided by NECTEC. For the companies that are not interested, increase in marketing effort can increase their knowledge in this field and might be interested in the future.

There are five main categories of the reasons that made the companies not interested in speech technologies:

1. Not related or suitable to their business
2. No application plan
3. No demand from customer
4. High investment cost
5. Other reasons not mentioned

Other reasons not mentioned include:

- The platforms that speech technologies can be used with do not match with the company's platform such as Window Mobile.
- The companies still do not have enough information about speech technologies.

According to Figure 33, the main reason is that technologies are not related or not suitable for their particular businesses with 58% of the total, which is more than half of all the reasons. Some companies said that they still do not know what applications that speech technologies can be applied to in their companies took over15%. Some said that there are no demand from their customers at 12%, and the reason that speech technologies include high investment cost at 7%. Some other reasons not mentioned above made up 8%.

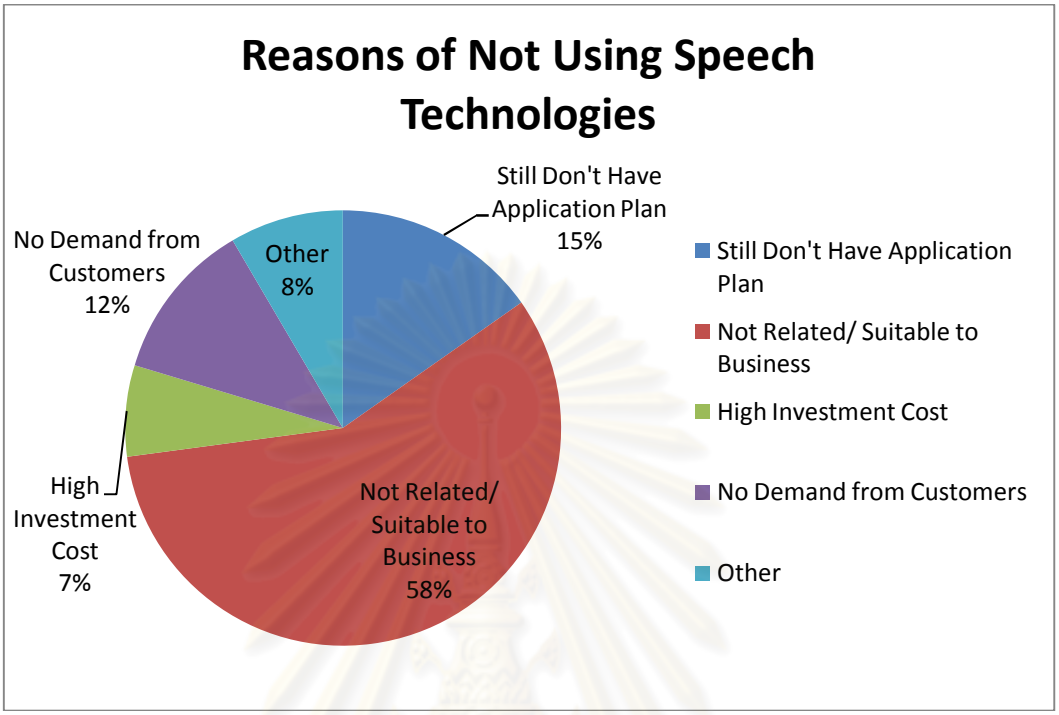


Figure 33 Reasons of Not Using Speech Technologies

According to Figure 34, as for the companies that are already aware of and understand speech technologies but still are not interested, the reason that the technologies are not related or not suitable to their companies made up over 69% of all reasons, which is very high percentage. The reason that the investment cost is high made up 15% of all given reasons. The reason that the companies still do not have the application plan for speech technologies made up 8%. Other reasons not mentioned above made up 8%.

It can be seen that the main reason that companies are not interested in using speech technologies is because they perceive the technologies as unrelated to their businesses. However, if more marketing and technology transfer by NECTEC to the businesses increased with supports in suggestions of application plans, customers might change their perception. As well as the companies that still do not know what type of applications that speech technologies can be used in their companies, help and support from NECTEC might increase the customer base. For the reason that

companies think that investment on speech technologies required high cost, the solutions might include negotiations with NECTEC for the suitable licensing plan, and support their decisions by trustworthy researches that support the investment on speech technologies.

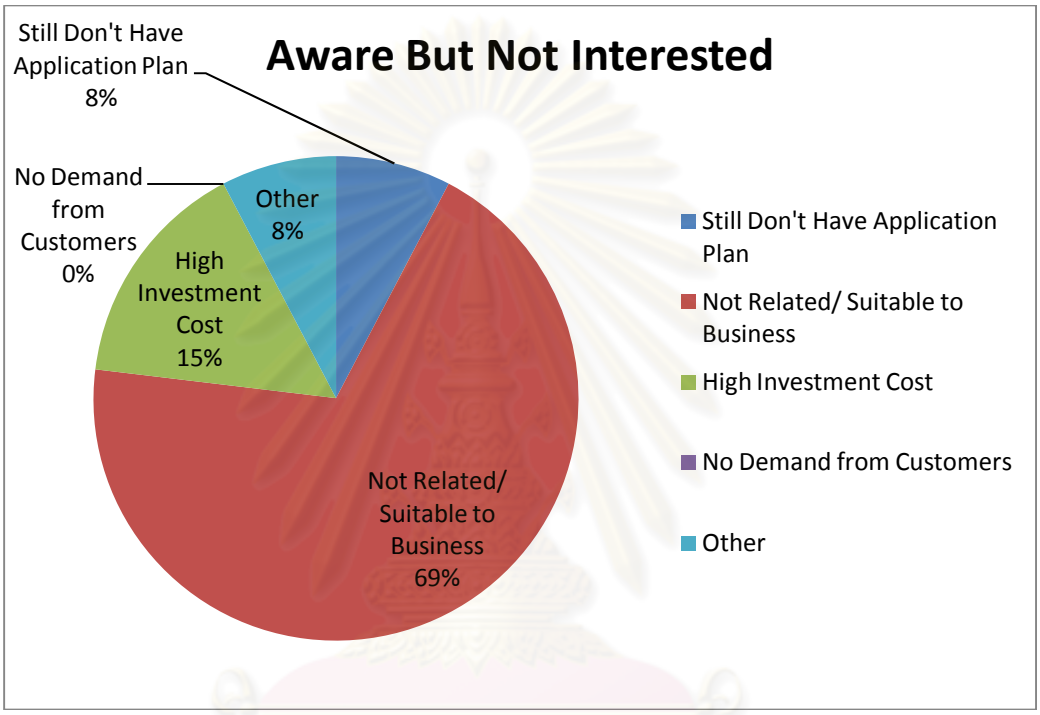


Figure 34 Aware but Not Interested

4.6.6 Opinions and Problems of Speech Technologies

This section will summarize about the opinions and perceptions that companies have toward speech technologies.

Direction of Speech Technologies in Thailand

The direction of speech technologies in Thailand from the companies' point of views can be categorized into four categories:

1. Increasing Sharply
2. Increasing Slowly
3. Constant

4. Decreasing

According to Figure 35, half of total respondents (50%) believed that the use of speech technologies will be increasing slowly in the several upcoming years. The main reason is that the technologies should increase along with the world's growth rate, but they are still very new in Thailand, therefore, the knowledge and technology transfer are still required. While nearly half (42%) believed that the speech technologies will be increasing their presence in several years with sharp growth rate, and the main reason is that the businesses in Thailand start to follow the world's trend and more companies see the potential of speech technologies.

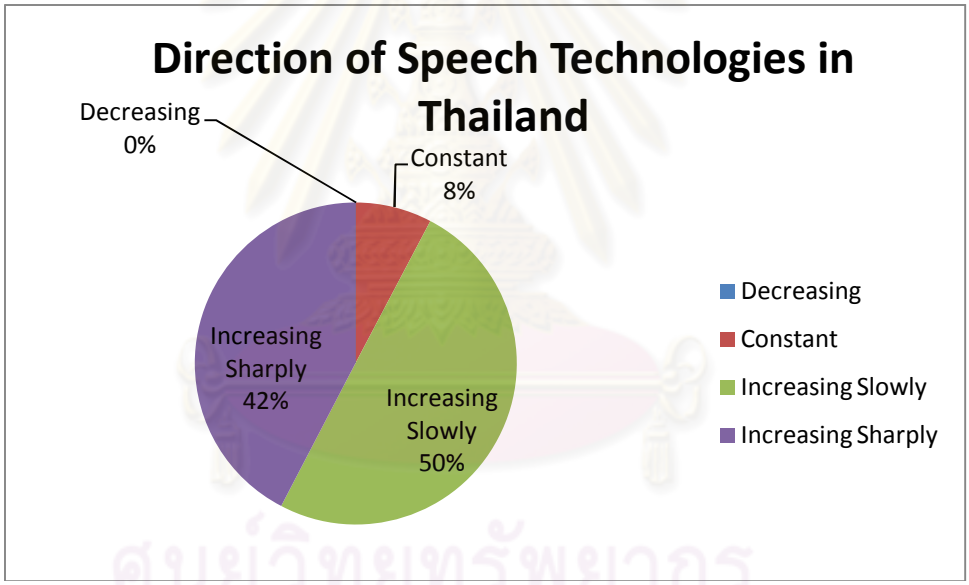


Figure 35 Direction of Speech Technologies in Thailand

Expectations of Speech Technologies

The expectation of speech technologies that the companies have can be categorized into four main categories:

1. Speech technologies will be widely used in business
2. Higher efficiency and reliability of the technologies
3. More speech applications introduced in Thailand

4. Other expectations not mentioned

Other expectations not mentioned include:

- More support from government
- More trial use from NECTEC
- Easy for developers to develop into applications

According to Figure 36, the main expectation of speech technologies that the companies have, which made up 49% of the total respondents, is that the speech technologies will be widely used in Thailand's business environment. The expectation of higher technical efficiency and reliability of speech technologies made up 21% of total respondents. The reason that more applications of speech technologies will be introduced and marketed made up 20% of all expectations. Other expectations not mentioned above made up 10% of the total. This can be implied that most companies are not quite sure about the trend of speech technologies in Thailand, which makes them uncertain about implementing the technologies.

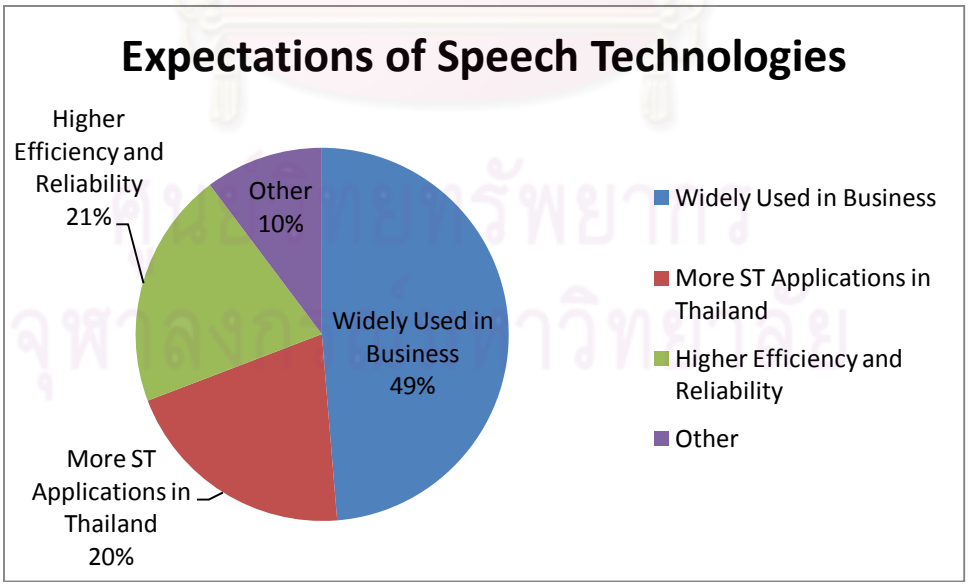


Figure 36 Expectations of Speech Technologies

4.6.7 Technical Problems of Speech Technologies

The technical problems that the companies encountered or are concerning that they could be happening mainly comprise of five main categories:

1. Accuracy of ASR technology
2. Efficiency and stability of speech technologies are not high enough
3. Quality of TTS technology is not up to standard
4. Complexity of implementation
5. Other technical problems not mentioned

Other technical problems not mentioned include:

- Some companies had experienced ASR technology many years ago (about 10 years ago) and perceived the ASR technology to be “speaker dependent”, and thus cannot be used with various customers. Therefore, NECTEC should inform and clarify this misunderstanding that the ASR technology has both speaker dependent and independent platform.
- The software’s memory requirement is not suitable for some platforms such as mobile platform. The software processing takes a considerably large size of memory requirement when used with mobile platform. NECTEC is now developing the speech software that will be suitable with mobile platform, which will be ready to be commercialized approximately by next year.
- Not enough technology transfer to other sectors
- Some companies feel that the ASR technology might be hard for the customer to use.

This part is conducted only with the companies that are aware of speech technologies. According to Figure 37, the accuracy of ASR is the main problem that companies are concerned of. Some believe that the accuracy of ASR is still considered

low, which is still not in the acceptable standard. This problem made up 39% of total problems that companies are aware of speech technologies. While the second most concerned problem is the efficiency and stability of speech technologies in Thailand. The companies are concerned that the technologies are not ready to be commercialized in Thailand due to the stability of the technologies, which made up 22% of the total. The third reason is about the quality of the TTS technology, in which the companies still think that the quality of the synthesized voice is still not up to standard and they would rather use the pre-recorded voice. The complexity of the implementation is another main reason that the companies are concerned of, which accounts 10% of the total problems. Other problems not mentioned made up 16%.

From this analysis, NECTEC should increase the quality of both ASR and TTS technology (more on ASR), and the companies should also be informed about that. At this point, NECTEC should significantly improve the quality of ASR. As for TTS, the technology is acceptable but should still be improved to increase customer satisfaction. However, NECTEC should also increase the technical understandings about the technologies between NECTEC and the companies as well.

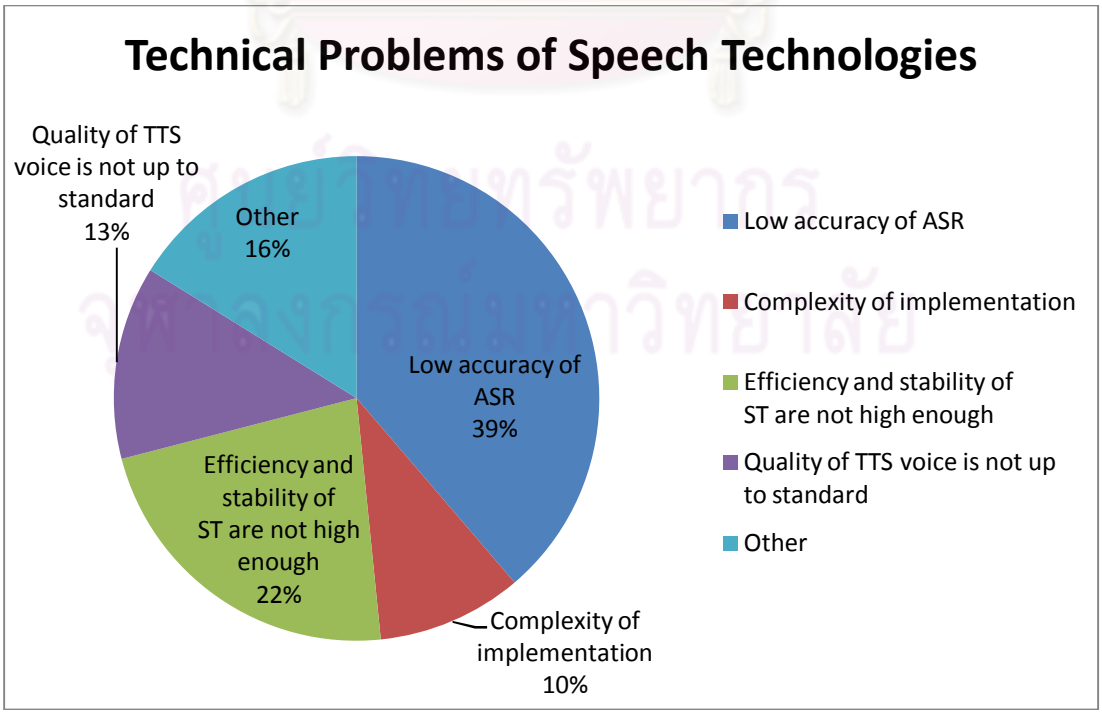


Figure 37 Technical Problems of Speech Technologies

According to Figure 38, as for the experienced user of speech technologies, the main concerns are similar, but some weights to the problems might be slightly different.

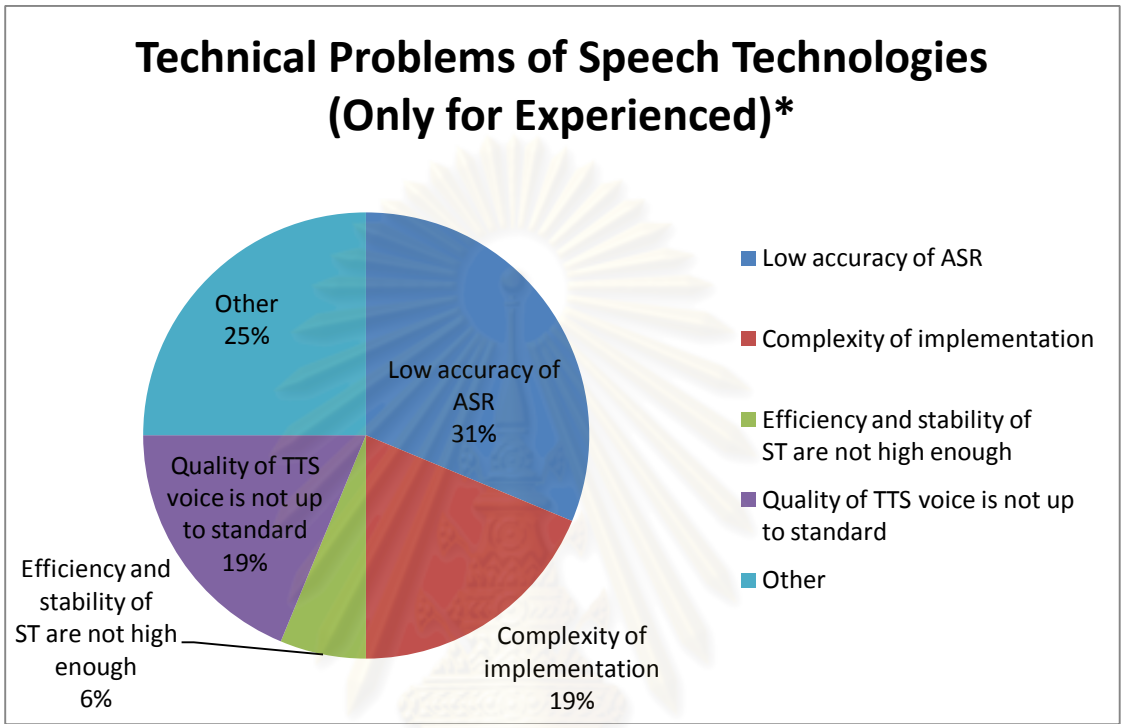


Figure 38 Technical Problems of Speech Technologies (Only for Experienced)

4.6.8 Business Problems of Speech Technologies

The business problems that the companies encountered or are concerning that they could be happening mainly comprise of four main categories:

1. Not enough knowledge / technology transfer and marketing
2. High investment cost, which leads to high selling point, and leads to low demand
3. Limited applications of speech technologies
4. Other business problems not mentioned

Other business problem not mentioned includes:

- Some other new technologies that are easier to implement or have lower cost might outperform speech technologies (substitutes).

This part is conducted only with the companies that are aware of speech technologies. According to Figure 39, the main concern of the companies is that the speech technologies are not well known in Thailand's business environment. More knowledge / technology transfer and marketing are needed in order to inform the users and the key players of the businesses. This accounts for 44% of all business problems concerned. The investment cost is another concern that companies have. The software houses who develop applications feel that the investment cost of speech technologies is high, which leads to high selling price of the applications and solutions and therefore, leads to low demand. This accounts for 33%. The third main reason that made up 17% is that the companies feel that the applications of speech technologies are limited to only some business area, and therefore, not suitable for many types of businesses. Other business problem made up 6% of the total.

It is clearly seen that increasing marketing efforts are required since companies do not have enough information about the speech technologies. More direct and indirect marketing efforts must be conducted. NECTEC might approach the interested companies with direct marketing efforts and provide road shows to reach mass amount of customers. As for the cost of implementing, NECTEC should provide trustworthy researches that support the investment on speech technologies can be beneficial. Also, NECTEC can support the companies with more application plans that the companies might be interested in.

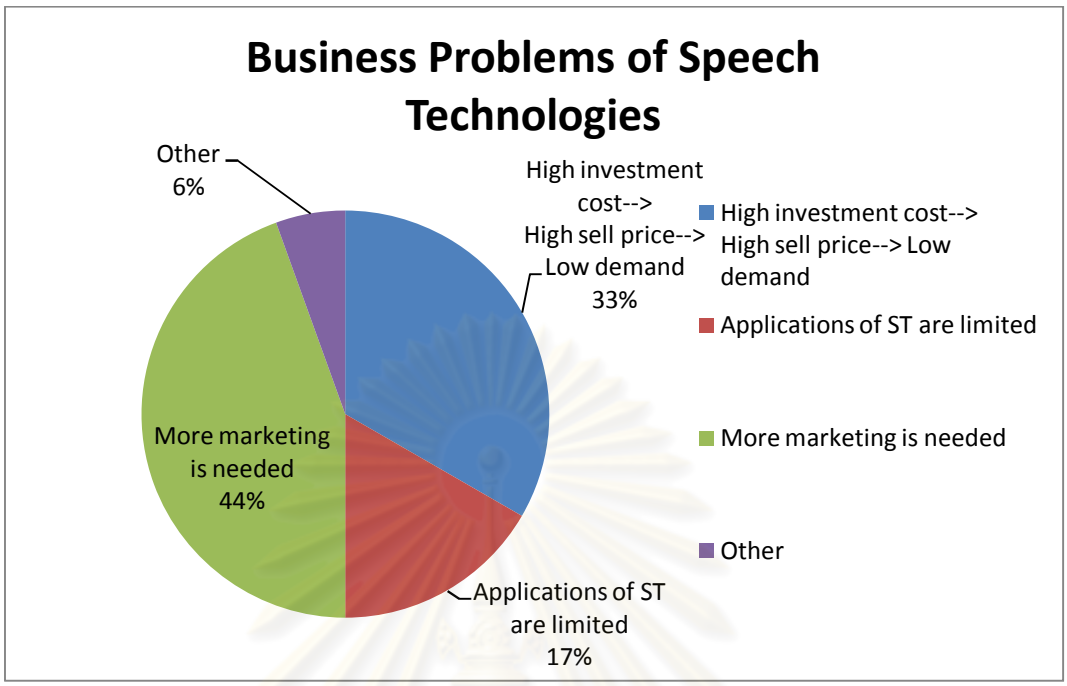


Figure 39 Business Problems of Speech Technologies

According to Figure 40, as for the experienced user of speech technologies, the main concerns are similar, but some weights to the problems might be slightly different.

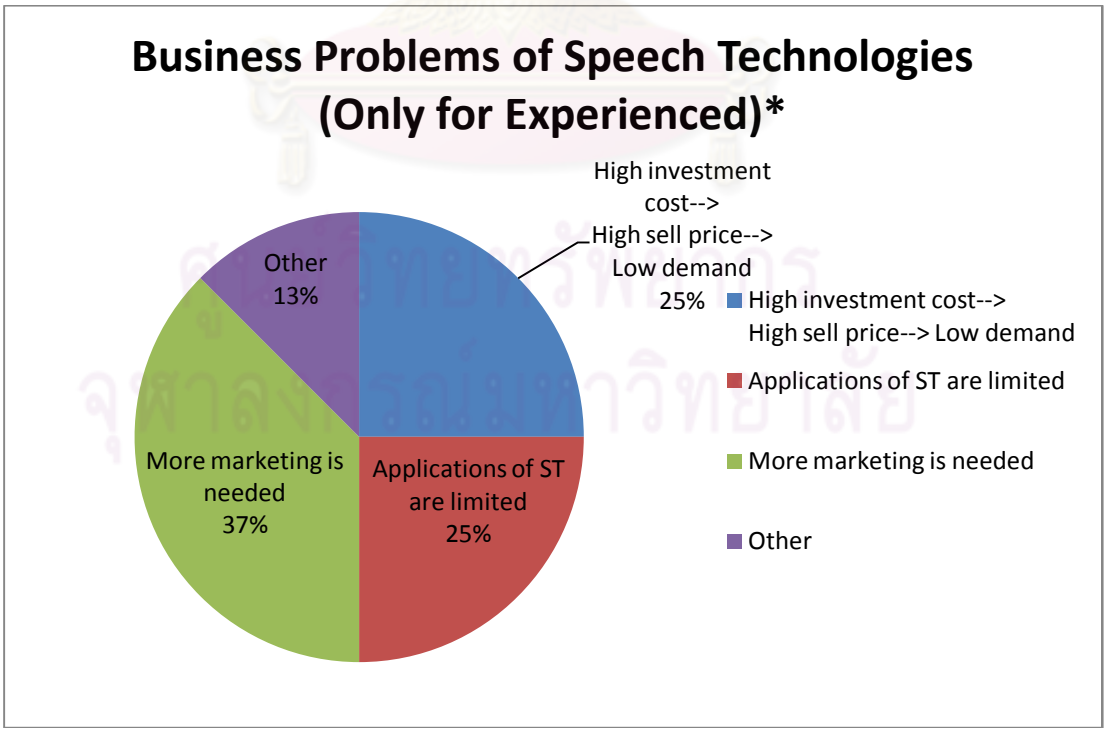


Figure 40 Business Problems of Speech Technologies (Only for Experienced)

4.6.9 Suggestions from the Market Research

Since the analysis suggested that the speech technology trend in Thailand should be increasing from the companies' point of view, NECTEC can use this as an opportunity to increase the customer base for their speech technology software. From the analysis above, the suggestions in for NECTEC to increase the customer base for speech technologies can be summarized as follows:

- Increase in awareness and experience of the technologies because understandings can lead to higher interest rate. The interest rate might be increased by increasing marketing effort and technology transfer.
- Provide trustworthy researches that support the investment on speech technologies can be beneficial.
- Place importance on the interested companies that already have plans in their mind and provide them with supports on implementations. Make sure that good customer relationship management is conducted with this group of customers.
- Provide interested companies that do not have application plans in their minds with plans and support.
- Negotiate and create suitable licensing plan with the interested customers.
- Improve the quality of the speech technologies and make sure to inform the customers about the improved quality. Also, clear understandings about the expectations of the customers should be created.

4.7 Market Segmentation and Targeting

The segmentation and targeting will be done base on the interview of the interested software houses, which are potential customers of NECTEC that might decide to use speech technologies for their applications. In this study, the market can be segmented by two main approaches: by the types of application and by the types of business the companies are in.

4.7.1 Market Segmentation by the Types of Application

According to Figure 41, from the total number of companies that are interested in speech applications and have plans to implement speech technologies with their businesses, there are six main market sectors segmented by the types of application:

1. Call Center – customer service channel by phones
2. Mobile Applications – applications for mobile phones using voice
3. Automation by Voice – using voice to control machines
4. Accessibility – increase accessibility among people by using voice to access data
5. E-book Reader – use TTS technology to read e-book files
6. Others

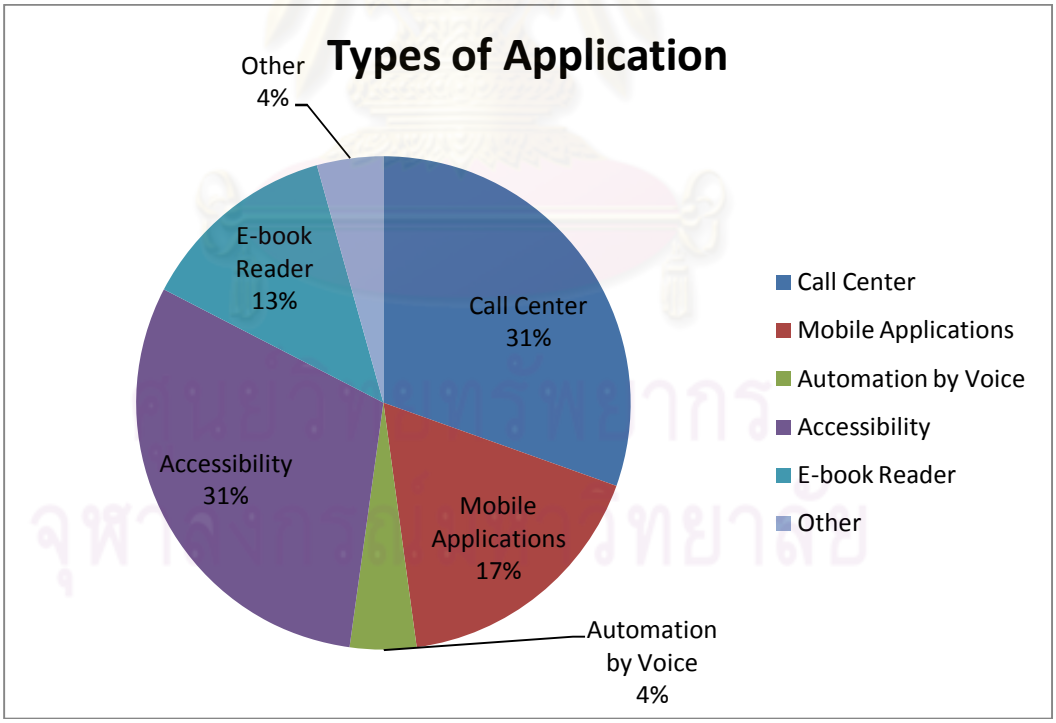


Figure 41 Types of Application

Call Center

The application segment with the highest interest rate of 31% is the call center applications. The interested companies plan to use TTS technology instead of recording voice, and use ASR technology as an option to touchtone menu. The suggested plans for call center applications are:

- Use TTS technology to provide information
- Use ASR technology as an option to touchtone menu
- Use ASR technology for voice search

Accessibility

Another application with the highest interest rate of 31% is the accessibility applications. Accessibility applications are applications that increase accessibility to data among people such as applications for handicapped or even for normal people to make their lives easier. The suggested plans for accessibility applications are:

- Use TTS technology for announcement in companies
- Voice navigated library system for the blind
- School's information system for parents (for example grading information that parents can have access to)
- Screen reader for web pages on internet
- Create response to user's click on web pages
- Use voice as another interaction channel for medical devices

Mobile Applications

The mobile application segment made up 17% of the total suggested applications. The ASR technology can be used to command the mobile applications and functions, and the TTS technology can be used to provide gimmicks for mobiles. The suggested plans for mobile applications are:

- Voice command for mobile device
- Voice search on mobile device
- Gimmicks for Android application

E-book Reader

Another increasing interest in technology gadget market is e-book reader applications. The e-book reader is the gadget used for reading e-book files as an option to read on laptop, which is more flexible because it is portable. TTS technology can be used with e-book reader to create gimmicks for the gadget in case the users want to listen to the books rather than reading by themselves. This section made up 13% of the total suggested plans. The suggested plans for mobile applications are:

- E-book reader
- Voice menu dictionary for restaurants

Automation by Voice

The automation by voice section made up 4% of the total number of plans. ASR technology can be used to automate machines or electrical devices as an alternative to pressing buttons. This application is suitable for working environment that workers are not hands-free such as in hospitals. The suggested plan for automation by voice is:

- Automation by voice for hospitals or other working environment that workers are not hands-free

Others

The last segment of the applications is the segments that are not mentioned above, which made up 4%. The suggested plan for other applications is:

- Talking dictionary – use TTS technology to read words and sentences and ASR technology to respond to user's input.

The types of application can be summarized in general as:

1. Application on PC
2. Mobile Application
3. Solution
4. Embedded Application (application embedded in other electrical devices other than computers)

According to Figure 42, the solution segment made up most of the total applications with 44%. The mobile application and embedded application made up 26% each. Finally, the application on PC made up the smallest portion with 4%.

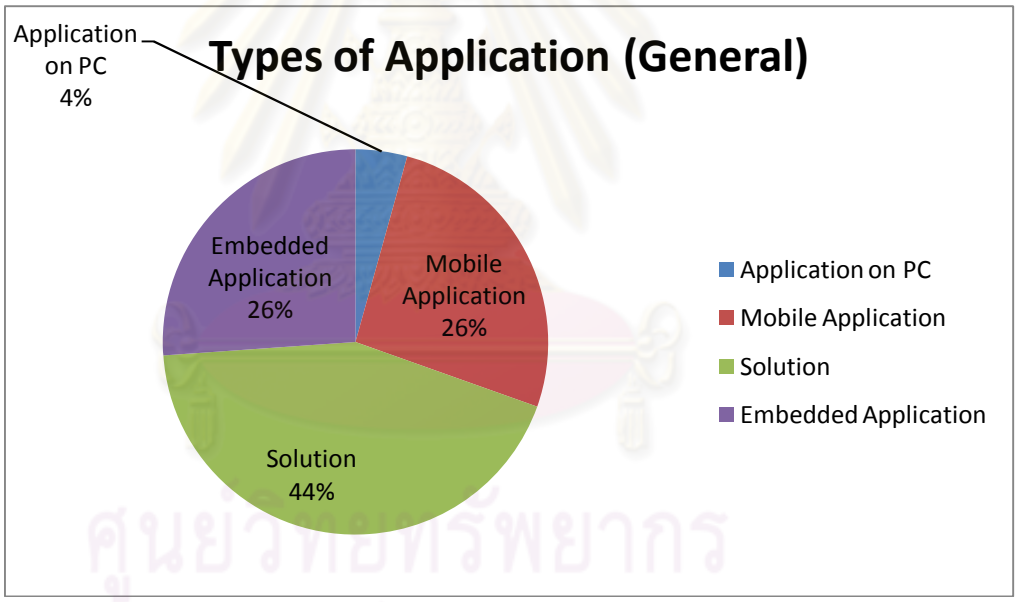


Figure 42 Type of Application (General)

4.7.2 Market Segmentation by the Types of Business

From the total number of companies that are interested in speech applications and have plans to implement speech technologies with their businesses, there are nine main types of businesses:

1. Telecommunications (mainly call center)
2. Office & Administrative Software

3. Account Software
4. Manufacturing Software
5. Retailing Software
6. Information Management Software
7. Mobile & Portable Devices
8. Media & Game
9. Other industries not mentioned

Figure 43 compares the interest rate among different types of business that are interested in speech technologies for example:

$$Interest\ Rate\ Among\ Different\ Types\ of\ Business = \frac{Number\ of\ Interested\ Telecommunications\ Companies}{Total\ Number\ of\ Interested\ Companies}$$

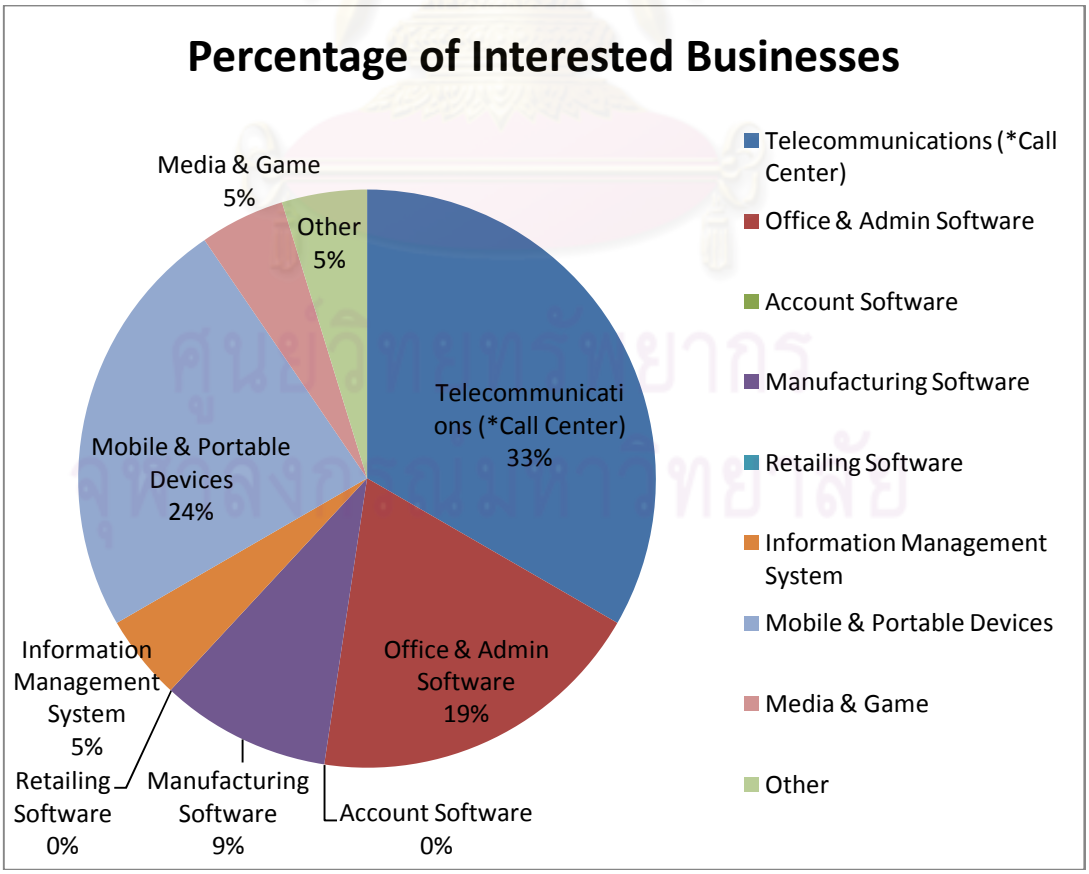


Figure 43 Interest Rate Among Different Types of Interested Business

Figure 44 compares the interest rate in each type of business for example:

$$Interest\ Rate\ in\ Each\ Type\ of\ Business = \frac{Number\ of\ Interested\ Telecommunications\ Companies}{Total\ Number\ of\ Telecommunications\ Companies\ Surveyed}$$

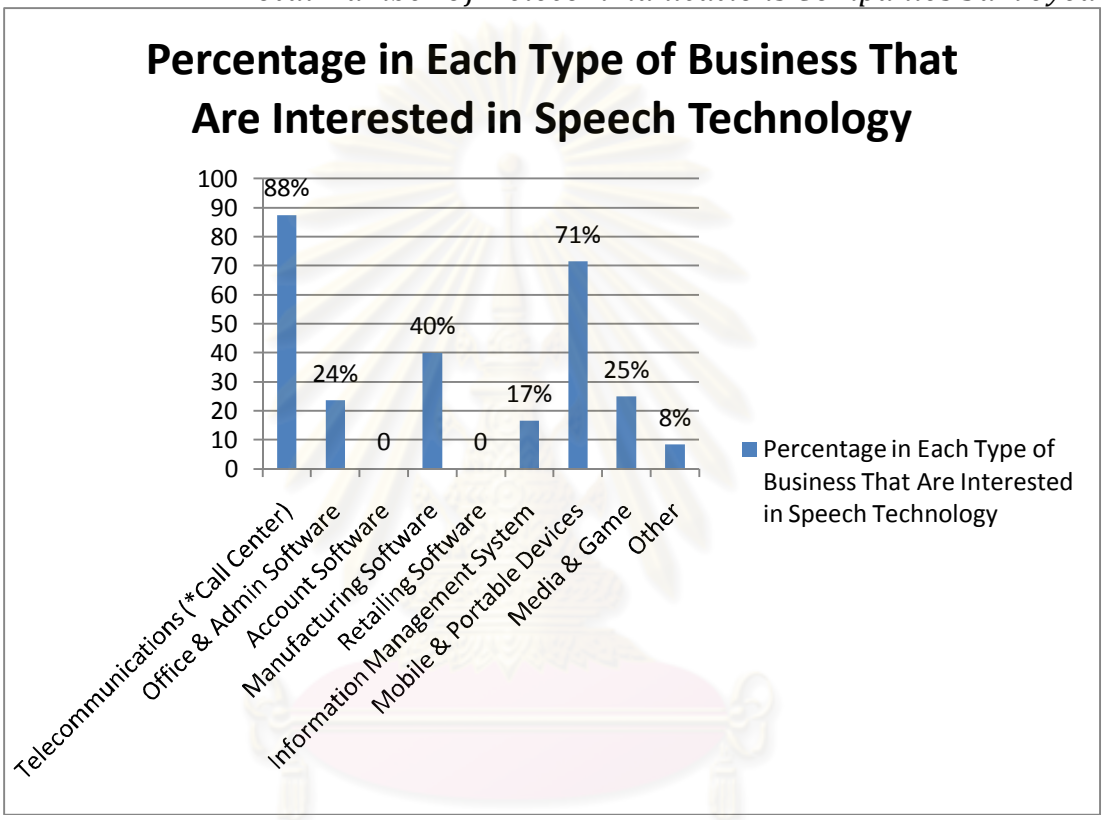


Figure 44 Interest Rate in Each Type of Business

*Telecommunications (*Call Center)*

The telecommunications business (mainly the call center business) has the highest interest rate of 33% comparing among other types of interested business (number of interested telecommunications companies to the total number of interested companies). From the previously market segmentation by types of application, the highest interest rate is also on the call center application. (Figure 43)

When comparing within the whole telecommunications sector surveyed (number of interested telecommunications companies to the total number of telecommunications companies surveyed), the interest rate is 88% which is very high. (Figure 44)

Office & Administrative Software

The office and administrative software has the second highest interest rate of 19% comparing among other types of interested business. (Figure 43) The applications for this business sector are mainly the accessibility applications, in which are planned to be used to add value.

However, if the interest rate is compared within only the office and administrative software sector, the interest rate is not high, only 24%. (Figure 44) This means that only a small percentage of this sector is interested in speech technologies. The reason that the interest rate is high when compared to other types of interested business is because of the abundance of the software in this sector in the business environment.

Mobile & Portable Devices

This sector also has the second highest interest rate of 24% comparing among other types of interested business. (Figure 43) This business segment corresponds to the application segment of mobile applications and e-book reader.

When compared within the mobile and portable devices sector, the interest rate is considered high with 71%. (Figure 44) This means that a large percentage of the companies in this sector are interested in using speech technologies.

Manufacturing Software

Manufacturing sector has the interest rate when compared to other interested types of business of 9%. (Figure 43) One might think that this business sector would correspond to the application of automation by voice, but actually automation by voice would not suit the manufacturing environment due to the loud environment. The loud surrounding noise would not be a suitable environment to use ASR technology; therefore, the automation by voice would not suit this sector as one might think. The survey actually shows that the plan for manufacturing sector is to use speech technologies for accessibility applications. The plan is to use TTS technology for announcement in factories.

When compared within the manufacturing sector, the interest rate is 40%. (Figure 44) This means that some percentage of the companies in this sector is interested in using speech technologies.

Information Management System

This sector has the interest rate of 5% comparing among other types of interested business. (Figure 43) This business segment corresponds to the application segment of accessibility applications. The plan for this sector is to use voice to respond to user's input when using the software.

When compared within the information management sector, the interest rate is only 17%. (Figure 44) This means that only a small percentage of the companies in this sector are interested in using speech technologies.

Media & Game

This sector has the interest rate of 5% comparing among other types of interested business. (Figure 43) This business segment corresponds to the application segment of accessibility applications. The plan for this sector is to use voice to respond to users in game applications.

When compared within the media and game sector, the interest rate is 25%. (Figure 44) This means that some companies in this sector are interested in using speech technologies, but not a high number.

Others

Other sector such as medical sector has the interest rate of 5% comparing among other types of interested business. (Figure 43) This business segment corresponds to the application segment of automation by voice. The plan for this sector is to use voice to control medical devices.

When compared within the group, the interest rate is only 8%. (Figure 44) This means that only a small percentage of the companies in this sector are interested in using speech technologies.

Account Software and Retailing Software

The account and retailing business sector have no interest in speech technologies according to the survey.

4.7.3 Technologies that Companies are Interested

The companies are asked about which type among the three types of speech technologies that they are interested in:

1. Text-to-Speech (TTS)
2. Automatic Speech Recognition (ASR)
3. Automatic Speaker Verification (ASV)

According to Figure 45, most interested companies are interested in TTS and ASR technology with 54% and 43% interest rate respectively. Only 3% are interested in ASV technology.

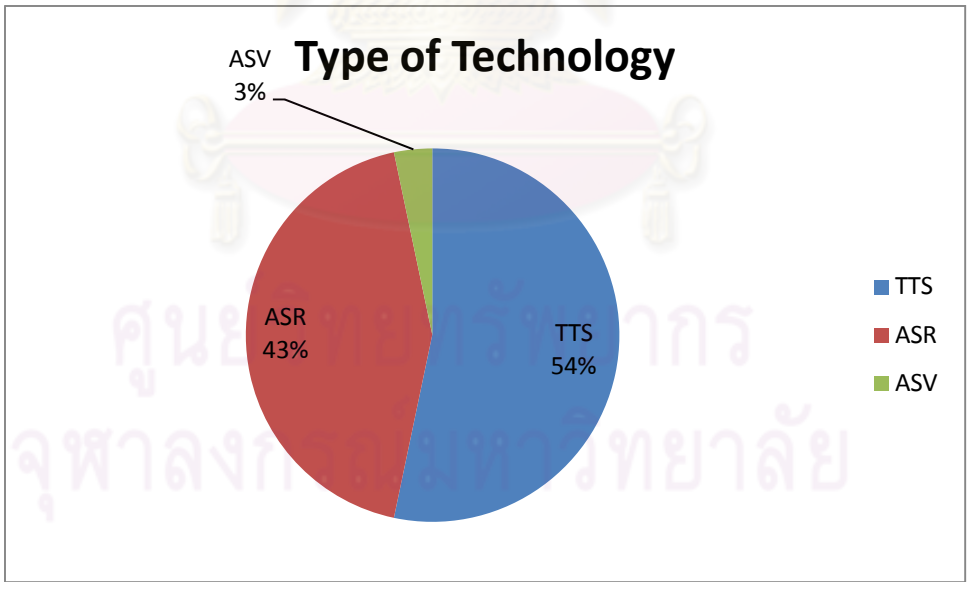


Figure 45 Type of Technology

4.7.4 Market of the Companies' Customers

According to Figure 46, the customer groups that the companies plan to use speech technologies with are mainly their existing customer group with 95% and plan to create new market with only 5%.

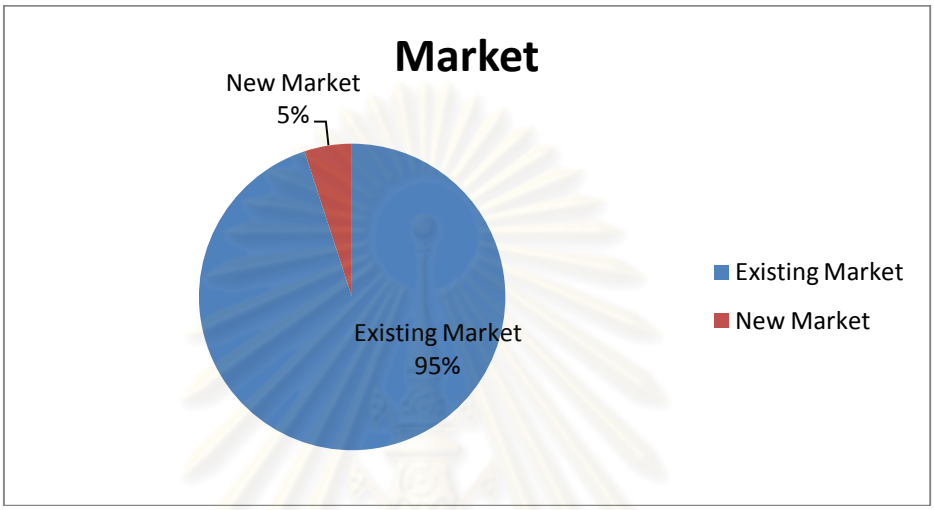


Figure 46 Market Aimed

4.7.5 Objective(s) of Using Speech Technologies

The companies were asked about their objective(s) of using speech technologies with their business. There are five main objectives:

1. Add value and create product differentiation
2. Increase efficiency of processes
3. Increase customer satisfaction
4. Create innovation and innovative image
5. Use as core application

According to Figure 47, from the survey, the main objective of using speech technologies is to add value to their existing product and create product differentiation with 41%. Because the speech technologies are new in Thailand, implementing it with the businesses will create innovative look for the companies. This objective made up 23% of the total objectives. Some companies believe that by using speech technology, customers will be more satisfied, and this is another main objective with 21%. Some

companies see the speech technologies as a tool to increase the processes' efficiency with 15%. However, none of the interested companies surveyed are interested to use the speech technologies as a core application.

It can be seen that almost all the companies plan to use speech technologies with their existing customers mainly to create the value-added functions that will differentiate their products from others. By implementing speech technologies the companies feel that it can increase customer satisfaction, create innovative image for the companies, and increase the efficiency of the current system.

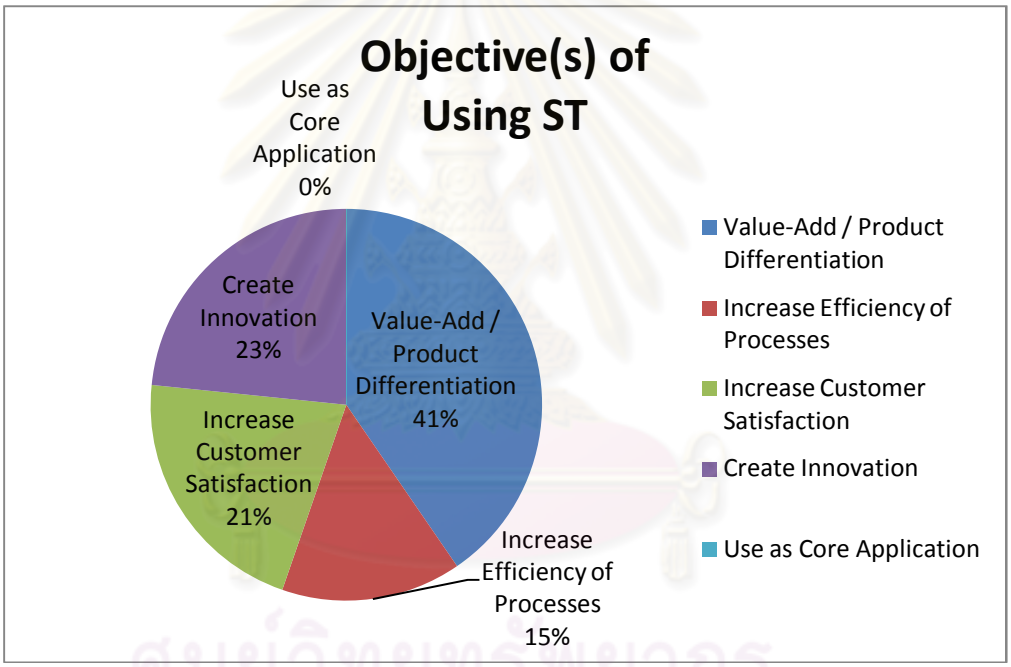


Figure 47 Objective(s) of Using Speech Technologies

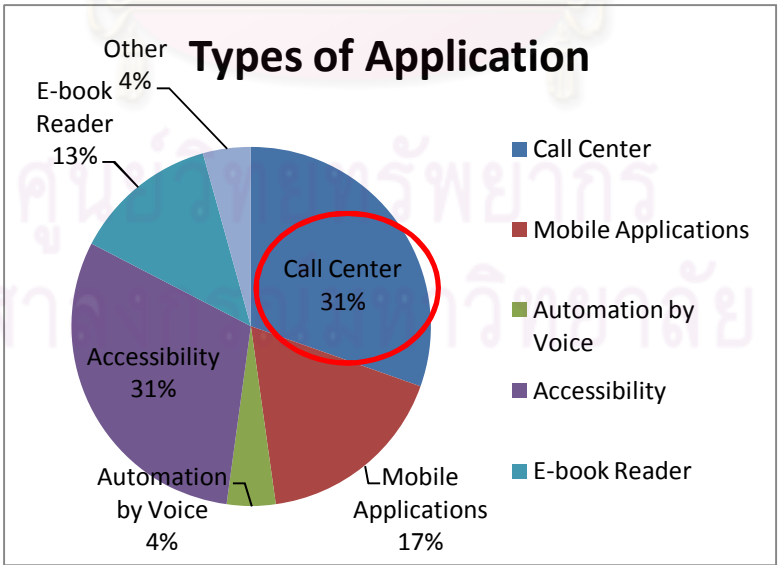
4.7.6 Market Targeting

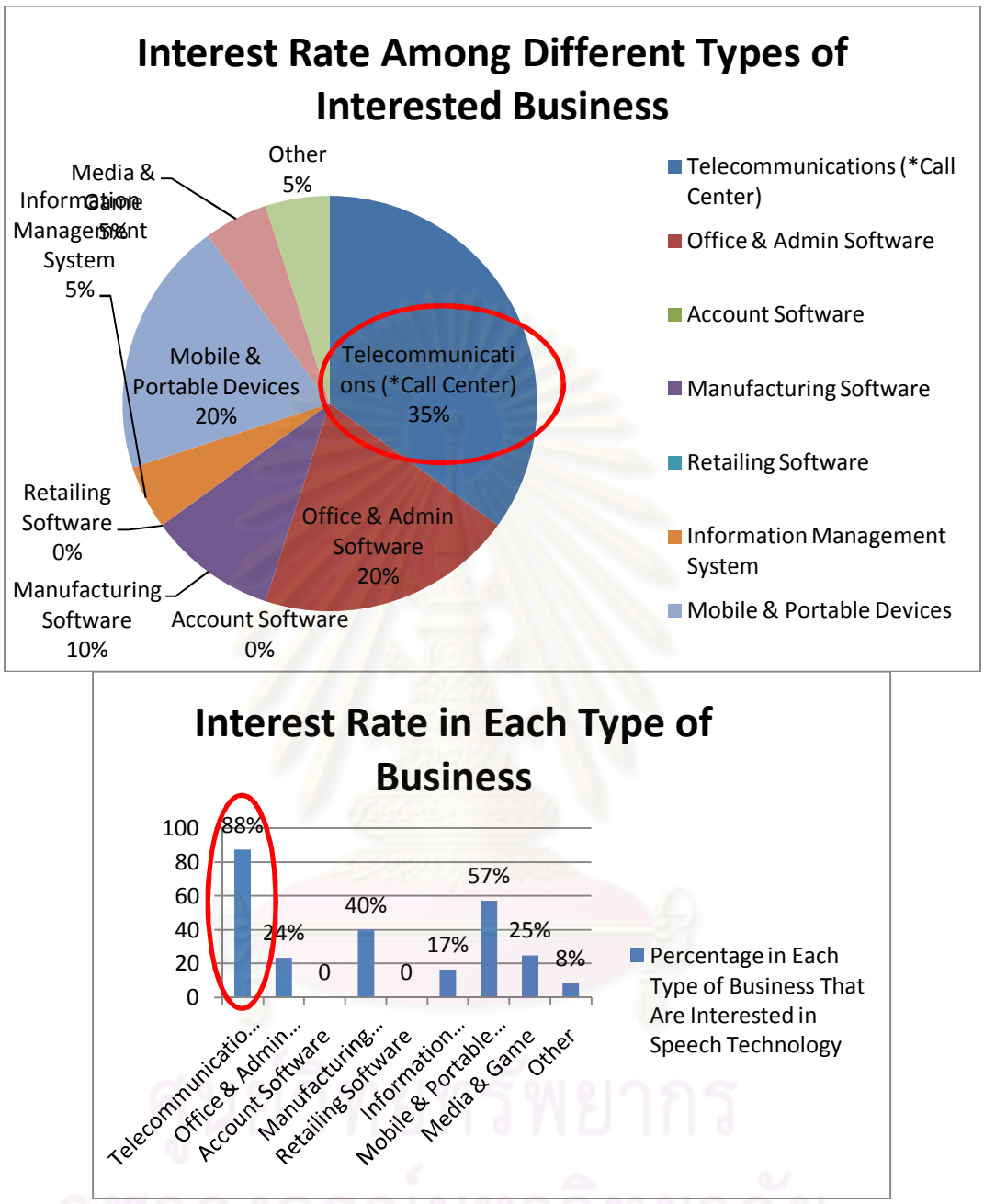
Current Target: Call Center Application

From the market segmentation and analysis above, the conclusion of the target market segment can be drawn as follows. The most prospective market segment for speech technologies in Thailand at present is the call center solution. As it can be seen that the market segmentation by the application suggested that the call center application has the highest interest rate along with accessibility application. Also when

looking at the market segment by the types of business, the interest rate is the highest in the telecommunication business (mainly call center businesses). Other than the demand from the call center business other reasons that support the targeting of call center business in the first stage are:

1. The nature of the call center business that speech is a relevant part of the business.
2. The attractive growth rate of call center business in Thailand: According to Frost & Sullivan Thailand call center market is forecasted to hit a year-on-year (YoY) growth of 18.4% in 2016. (Frost & Sullivan, cited in The Nation Newspaper, 2010)
3. Successful international use of speech technologies with call center business such as in the United States and the United Kingdom.
4. Increasing competition in the call center business, which calls for the value-added functions, in which speech technologies can be the prospective choice.
5. The call center software group is easier to target since they are quite similar in nature among the group and can directly target at telecommunications sector.

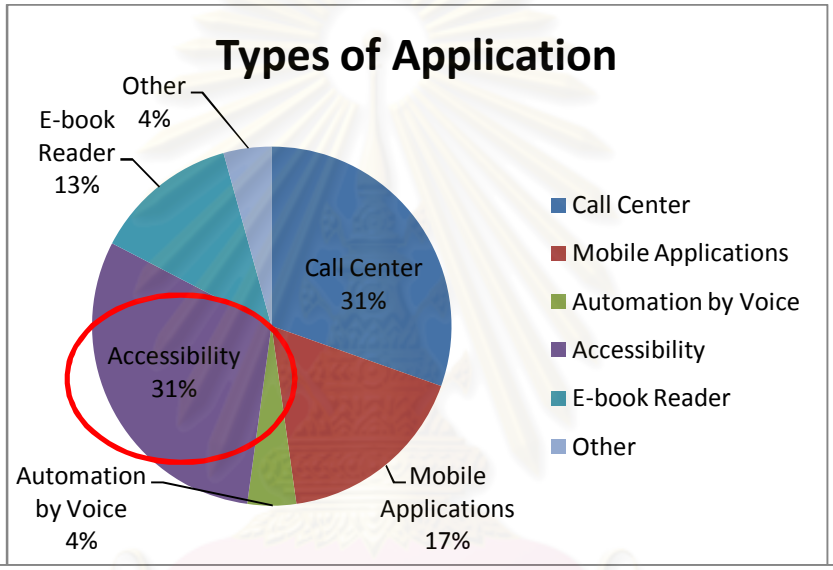


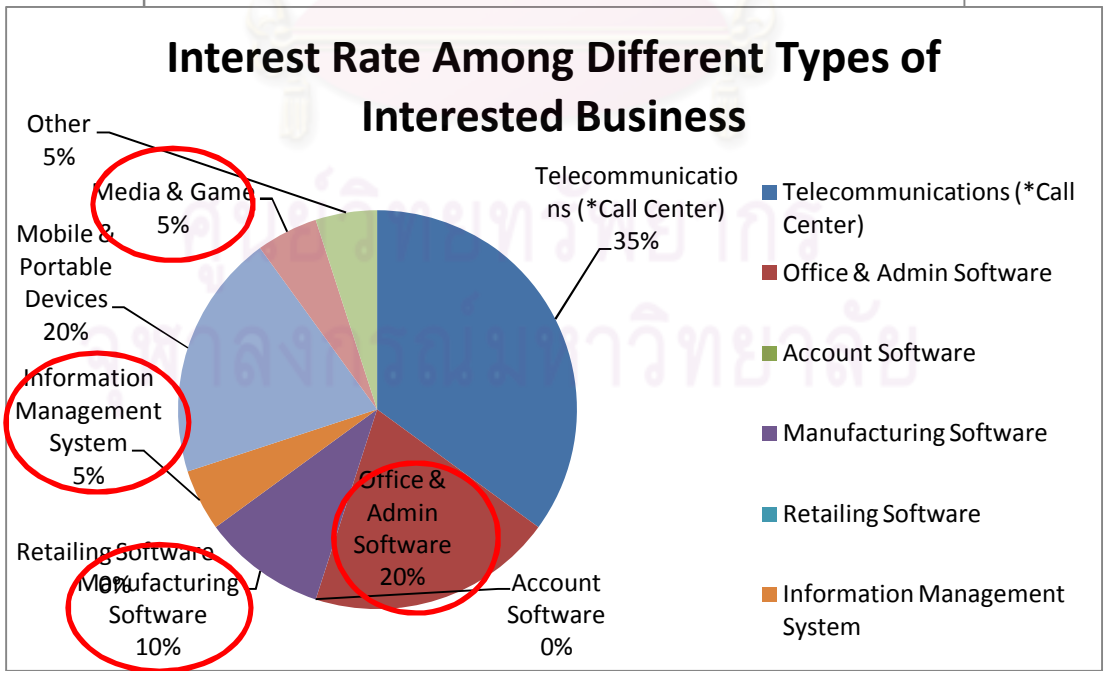


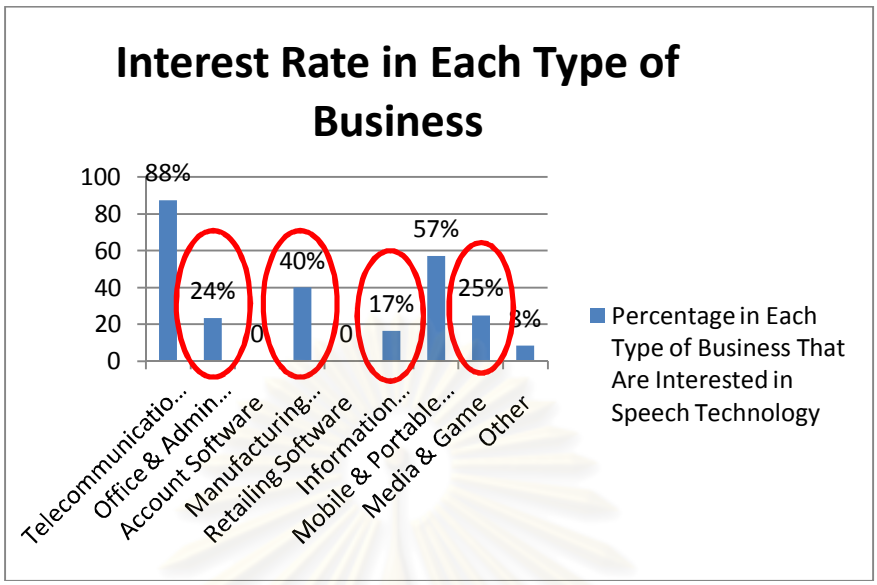
Future Target: Accessibility Application

The second most interesting application is accessibility applications with the highest interest rate along with the call center application. However, the accessibility applications are suitable to various types of business: office & administrative software, manufacturing software, information management system, media & game, etc. Since there are many types of business that are suitable for the accessibility application, the

customer base is large but hard to focus. It must also be noted that the interest rates comparing within each type of business are low for the suggested types of business. In conclusion, accessibility application has very wide ranging solutions and can be used in a wide variety of businesses, but the drawback is that it would be hard to focus on a particular group of customer. In the future with more technology and knowledge transfer, the interest rate could be increased by giving the potential customers more information.

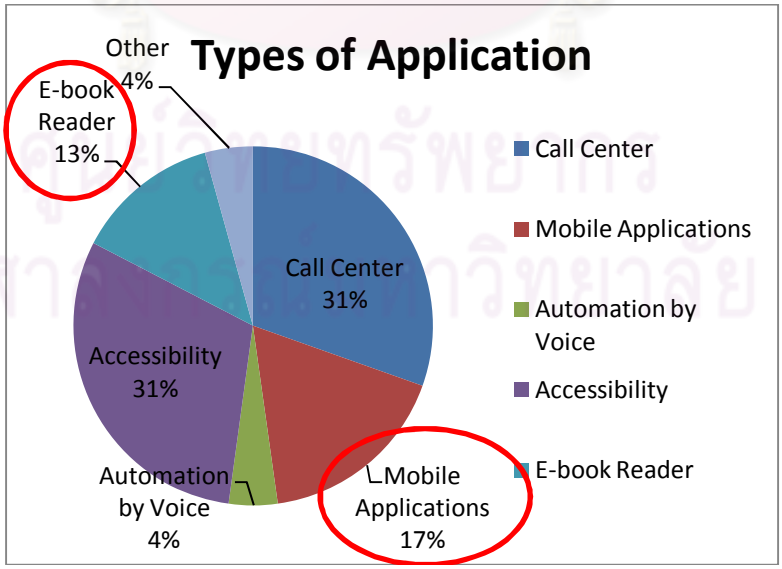


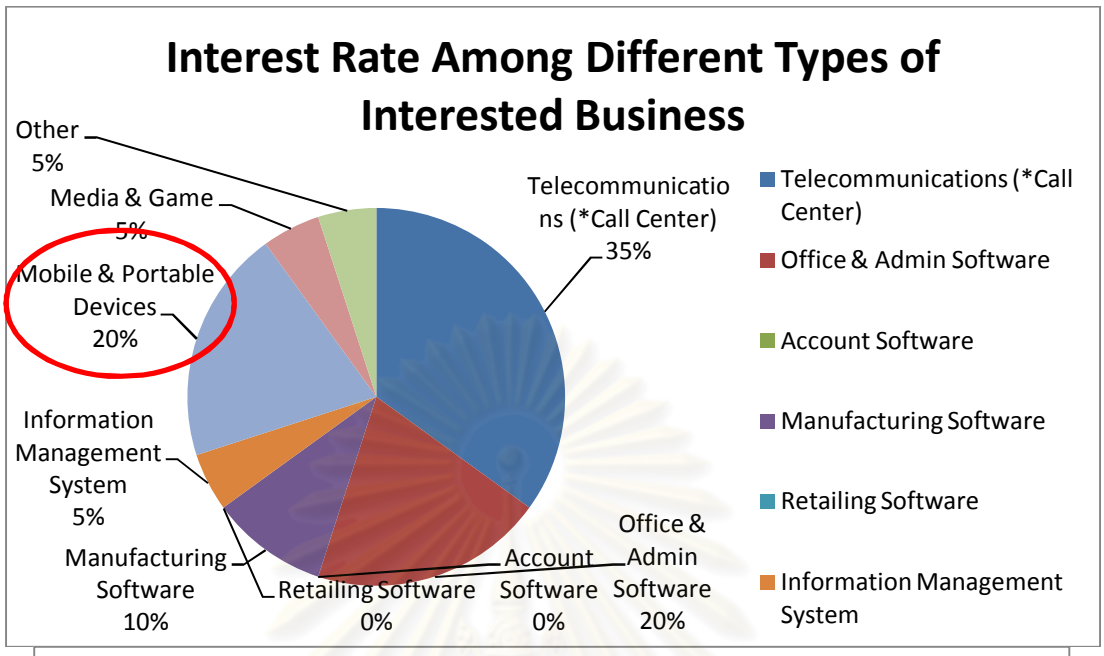


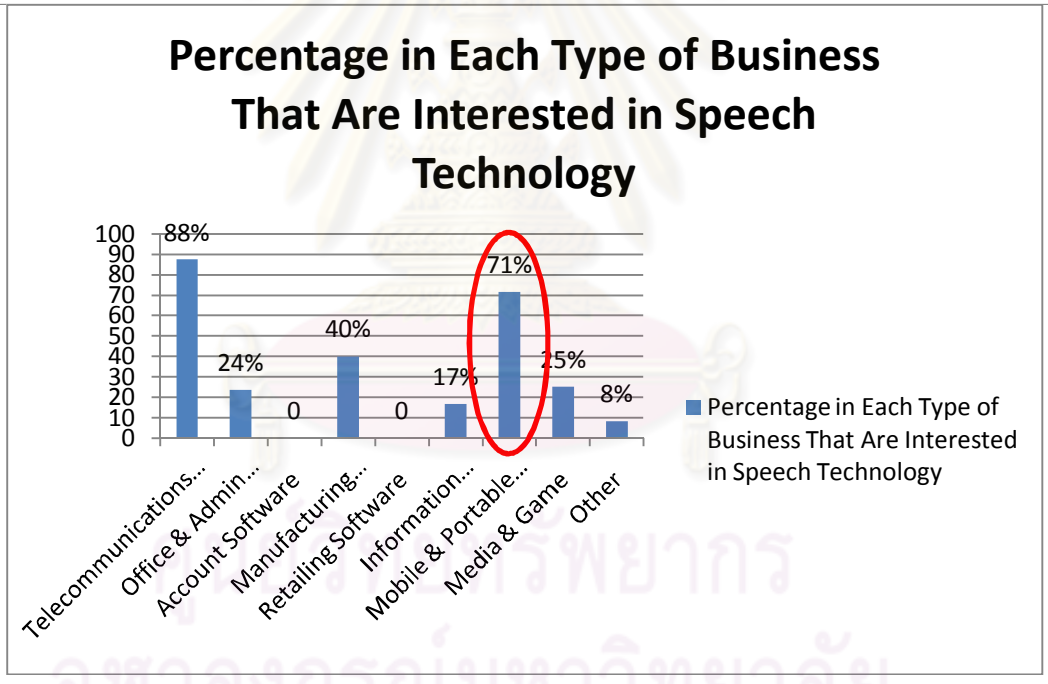


Future Target: Mobile Applications and E-book Reader

Another interesting application of speech technologies are the mobile applications and e-book reader due to their high interest rate. However, due to the technological readiness of the speech technologies of NECTEC that are not suitable with mobile platform at the present, the mobile applications can be an interesting future market.




Interest Rate Among Different Types of Interested Business
Other 5%
Media & Game 5%
Mobile & Portable Devices 20%
Information Management System 5%
Manufacturing Software 10%
Retailing Software 0%
Account Software 0%
Office & Admin Software 20%
Telecommunications (*Call Center) 35%
Telecommunications (*Call Center)
Office & Admin Software
Account Software
Manufacturing Software
Retailing Software
Information Management System


Percentage in Each Type of Business That Are Interested in Speech Technology
100
90
80
70
60
50
40
30
20
10
0
88%
24%
0
40%
0
17%
71%
25%
8%
Telecommunications...
Office & Admin...
Account Software
Manufacturing...
Retailing Software
Information...
Mobile & Portable...
Media & Game
Other
Percentage in Each Type of Business That Are Interested in Speech Technology

Chapter V
Targeting the Call Center Business

5.1 Background of Call Center Business

5.1.1 Overview

Nowadays, telephone is still the most heavily used of communication channel due to the convenience and the reasonable cost. Real Bergevin and others stated that “To the customer, the call center is the voice of the company.” (Bergevin et al., 2010) By creating customer satisfaction when a customer calls, the loyalty can be created and initiate repeating service afterward. According to Sorawit Buasri, the objective of a call center is to increase the service efficiency to the customers via telephone with the focus on improving service by the selected human resources and computer technology. (Buasri, 2009) According to Real Bergevin and others, in the past a call center was viewed as a “cost center”, but nowadays it is rather a “profit center” that can increase the competitive advantages and customer relationship for a company. (Bergevin et al., 2010) A call center might provide information about the products or services, take orders, give technical supports, receive customers’ complaints, etc.

There are three main categories of call centers: (Bergevin et al., 2010)

1. Inbound- The customers initiate the calls to the company’s call center for inquiries or assistance.
2. Outbound- The agents initiate the calls to the customers for telemarketing campaign, send out notifications, or create customer relationship.
3. Blended- Some call centers are both inbound and outbound.

According to Real Bergevin and others, four main goals of a good call center are as follows: (Bergevin et al., 2010)

Figure 48 Interdependent Business Goals (Bergevin et al., 2010)

From “Call Centers for Dummies: 2nd Edition”, 2010

5.1.2 Call Center Technologies

The important component of a call center is the human agents or operators. However, nowadays, call centers are trying to manage the human resource efficiently by introducing different kinds of new technologies. The effective blend of human resource, processes, and technologies will leads to the optimum level of the business goals suggested above. The important call center technologies used in Thailand are: Automatic Call Distribution (ACD), Computer Telephony Integration (CTI), and Interactive Voice Response (IVR).

Automatic Call Distribution (ACD)

Automatic Call Distribution (ACD) system manages thousands of incoming calls and routes to the agents. Each agent will receive approximately equal number of incoming calls. The incoming calls will be queued up, and the longest-waiting calls will be transferred to the available agents first. The queuing system is a key part in the call

center's efficiency because it is related to the waiting time that the customers have to wait. (Bergevin et al., 2010)

Computer Telephony Integration (CTI)

According to Sorawit Buasri, Computer Telephony Integration (CTI) integrates the telephony system with computer network. The CTI system will search the database system and integrates the incoming calls with the information from the database. The information about the customer will be popped up on the computer screen of the agent that takes the call. Also, CTI system keeps track of the incoming data in the database as well. The CTI system increases the efficiency of the call center operations and customer relationship management significantly. (Buasri, 2009)

Interactive Voice Response (IVR)

Interactive Voice Response (IVR) uses automated voice to response to the customers. The system allows the call center to work 24/7 in cases such as answering commonly asked questions, get information, or complete transactions. According to Sorawit Buasri, financial and banking institutions are one of the most important IVR user group to provide services such as providing account balance information, transferring, providing general information such as currency exchange rates, cancel the ATM cards, etc. (Buasri, 2009) According to Real Bergevin and others, IVR system is very cost-effective, and the cost of a service provided by IVR can be less than one fifth of the cost using the human operators. (Bergevin et al., 2010)

In the past, IVR system enables the self-service system for customers either by keyboard input (commonly known as "touchtone system") or speech recognition. In Thailand, most IVR system uses only keyboard input and responds with pre-recorded audio, but in reality some data cannot be input by the keyboard such as the address, and some data cannot be pre-recorded. The TTS and ASR technology are increasingly applied to the IVR system in other countries. The TTS technology can replace the old pre-recorded voice, and the ASR technology can replace the keypad input.

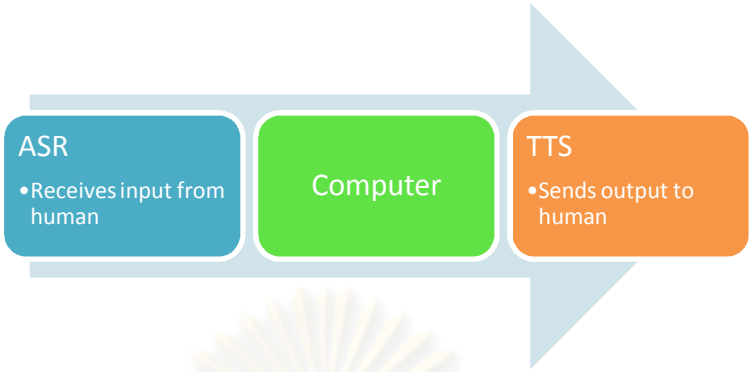


Figure 49 IVR System

The call center operations can be briefly summarized in the following figure:

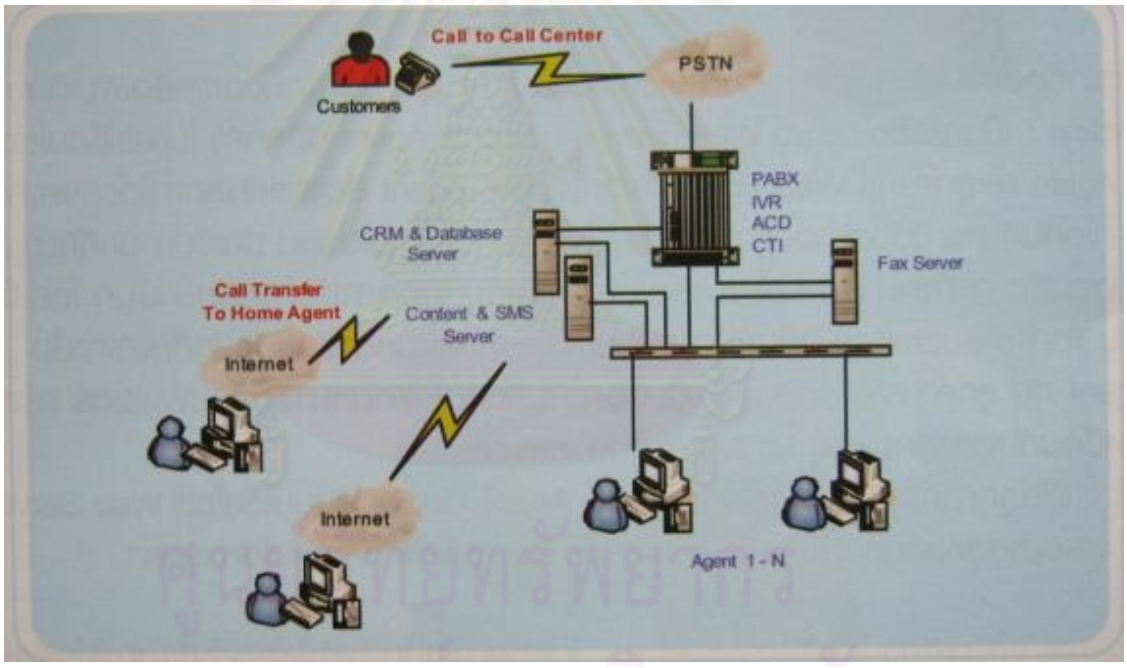


Figure 50 Figure Call Center Operations (Buasri, 2009) From “Thailand Contact Center 2009”, 2009

5.2 Reasons to Employ Speech Technologies in Call Centers

There are several reasons to why speech technologies should be employed in call centers:

- Improve customer satisfaction by convenience ➔ It is obviously easier for a person to say out their request or command than by pressing keypads. As Frank

Sherlock, the senior vice president and managing director of Intervoice International, suggested that “The human voice is the most natural interface in the world.” (Frank Sherlock, cited in Voice Compass: International 2008/2009: Speech Goes Mainstream by Artelt, 2008) Therefore, by using speech technologies, users will be more convenient to just speak their request. This allows users to just say what they want instead of taking the phone off their ears to press the keypads every time they need to request something.

- Bypass the menu ➔ Long menu can be bypassed by employing speech technologies. With the existing touchtone system, the users have to wait and listen for the ongoing lists of menu and remember the number they have to press, for example, “Press 1 to download music, press 2 to listen to the information, press 3…”. However, with speech technologies, the menu can be shortened, for example “Say download music, information, or…” or even with the more sophisticated system the user can just say their request in any way they want. The users do not have to remember the long list of numbers that they have to press anymore.

- Increase the range of data input and output ➔ By implementing speech technologies, users can obtain all kind of information by phones. In the past, some information might be hard to input by keypads and get the dynamic output such as the names, address, etc. According to Detlev Artelt, almost all kind of information can be obtained by phones with speech technologies. (Artelt, 2008)

- Automate repetitive inquiries and transactions ➔ Some repetitive tasks such as general asked questions, simple inquiries can be carried out by automation of speech technologies without requiring human agents.

- Increase efficiency and decrease operational costs ➔ The efficiency can be increased by implementing speech technology. For TTS technology, the pre-recordings of the information that are dynamic and change frequently are not required anymore. The time and cost of pre-recordings are eliminated. Simple and repetitive tasks are not required to be carried out by the human agents anymore, which lead to the shorter waiting time for the users. Therefore, the total time for a user in the system will decrease, and the system will be able to support more users.

- Optimize human resource and their talents ➔ By implementing automated speech technology system does not mean that the human agents are made redundant but to allow them to manage the more complex and important tasks. (Artelt, 2008) The human agents will have more chances of dealing with the more complex tasks rather than repetitive tasks, which will improve the pride in their jobs.

5.3 Value Chain of Speech Technology in Call Center Business

The call center business in Thailand can be segmented into three main service sectors: outsourced call center service, call center system and software service, and telemarketing service. (Kasikorn Research Center, 2009) In this study, the focus is mainly on the call center software service.

1. Outsourced call center service- Provide the whole solutions and infrastructure of call centers with the agent service for companies to outsource call center function.
2. Call center system and software service- Provide the solutions of call centers and infrastructures to companies to do in-house call center function.
3. Telemarketing- Outbound call centers with the telemarketing function.

The value chain of speech technology from the basic technology provider (NECTEC) to the end users of the call center business in Thailand is as follows:

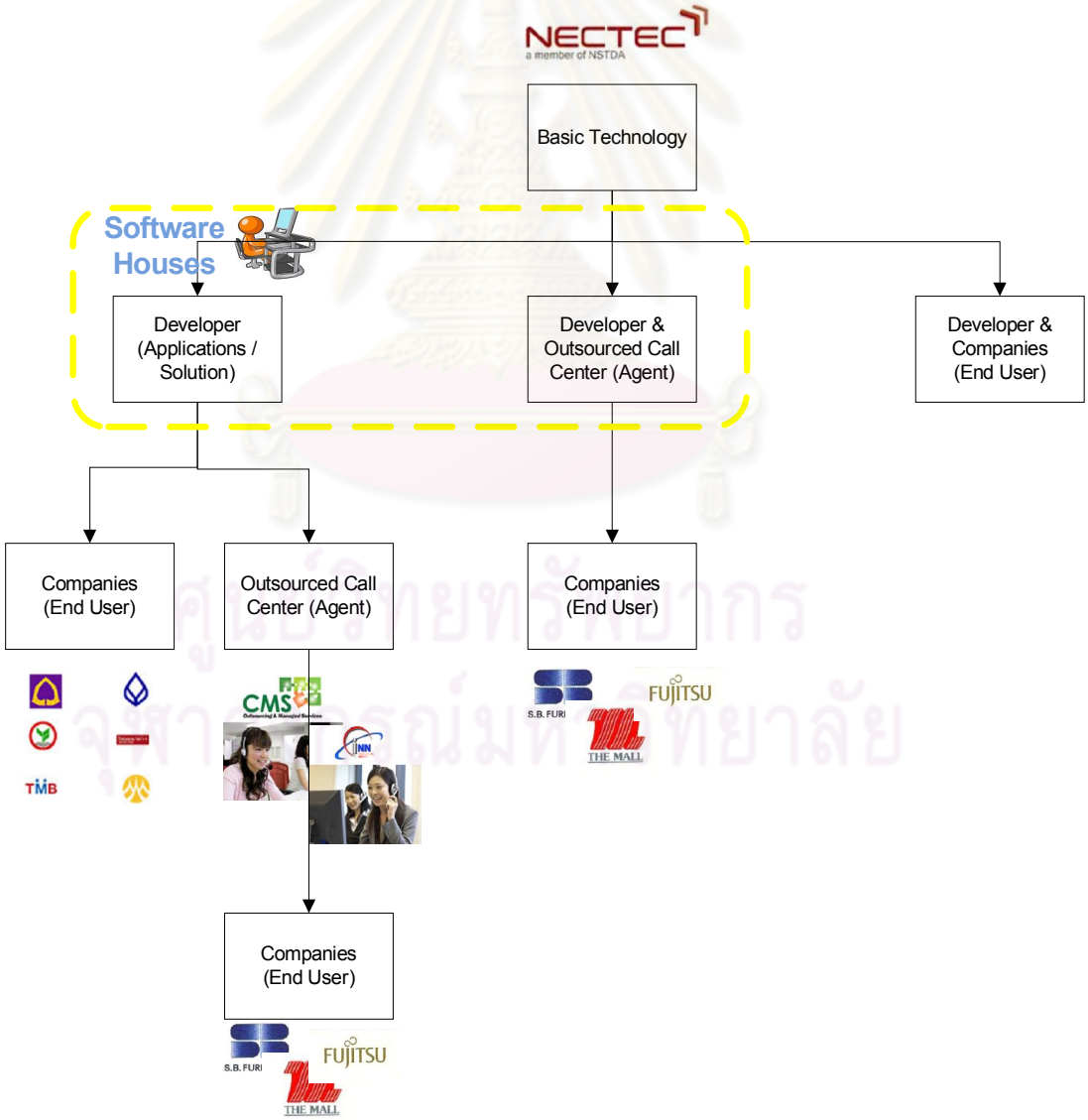


Figure 51 Value Chain of Speech Technology in Call Center Business

This diagram is based on the interview of the key players in the call center business in Thailand. Each participant in the diagram is explained as follows:

1. Basic Technology Provider- The basic technologies (TTS and ASR software) are developed by the basic technology provider, in this case is NECTEC. There are also other basic technology providers as stated in the previous section as well:

   TTS Providers: NECTEC, Chulalongkorn University, Private Companies
   ASR Providers: NECTEC, Chulalongkorn University, Private Companies

2. Developers (Software Houses) - VAJA 6.0 and iSpeech are the basic technology software developed by NECTEC, which is then licensed to the developers. The developers develop call center applications / solutions from the basic technologies.

3. Outsourced Call Center (Agent) - The third party outsourced call centers provide the call center service to the companies that want to outsource call center function. The outsourced call centers provide the companies with everything required to operate the call center function from the hardware, software, and the human operators (agents).

4. Companies (End Users) - The companies are the end users, in which the call center function can either be in-house or outsourced.

Each value chain is explained as follows, starting from the most left value chain:

1. The first and most abundant group of potential customers of NECTEC is the software houses. The software houses license the basic technologies and developed into solutions for the call centers. The software houses then sell the solutions to either the companies (end users) that do in-house call center, or sell to the outsourced call centers. The outsourced call centers then sell the solutions in combination with providing agent service to the companies (end users)

2. In some cases, the outsourced call center parties develop their own solutions without buying from other software houses. It can be said that they are both the developers and the outsourced call center service providers. The solutions and services are then bundled and sold to the companies (end users).

3. The last case is the companies (end users) buy the basic technology directly from NECTEC and develop their own call center solutions and do in-house call center all by themselves.

However, the most abundant group of potential customers is the software houses that are specialized in developing call center solutions. This study will mostly focus on the first and second group to customers (the software houses and the outsourced parties that develop their own software). The last group (the companies that develop their own software for the call center) is very rare.

5.4 Status of Call Center Business in Thailand

5.4.1 Call Center Business Growth and Revenue

According to Kasikorn Research Center, the total market worth for call center business in Thailand in the year 2008 was approximately 3,000-3,500 million THB (approximately US$90-110 million). (Kasikorn Research Center, 2009) Frost & Sullivan suggested that Thailand call center market is forecasted to hit a year-on-year (YoY) growth of 18.4% in 2016 and Thai sector is projected to reach 63,000 call center seats in the near future. (Frost & Sullivan, cited in The Nation Newspaper, 2010)

Total market worth for call center business in Thailand (2008)	3,000-3,500 million THB (approximately US$90-110 million)
Forecasted call center business year-on-year (YoY) growth in 2016	18.4%

Table 14 Call Center Business Growth and Revenue

(According to Kasikorn Research Center and Frost & Sullivan)

However, the annual growth of the call center industry in Thailand starts to slow down comparing to the past annual growth of approximately 20-30%. (Kasikorn Research Center, 2009) According to a key player interview of a call center software company, the current growth in the high-revenue market is not more than 10% because almost every large organization already owns a call center. Although the growth for the high-revenue market is slowing down, this will be compensated with revenues from small to medium size companies that still start to initiate or improve call center operations. The call center industry still has the room to grow and improve because companies are now placing a far greater importance on customer satisfaction, which good call center operations can contribute greatly to it.

According to Frost & Sullivan, the growth of the call center business in Thailand is closely linked to the domestic economy. The previous period of global economic crisis in combination with the political instability in Thailand led to the significant decrease in

the revenue, but is expected to recover in short to medium timeframe. (Frost & Sullivan, cited in The Nation Newspaper, 2010)

5.4.2 The Customers of the Call Center Business

Kasikorn Research Center suggested that the main group of call center end users is the business sectors at the rate of 70% and the leftover 30% are the government sectors. (Figure 52) (Kasikorn Research Center, 2009)

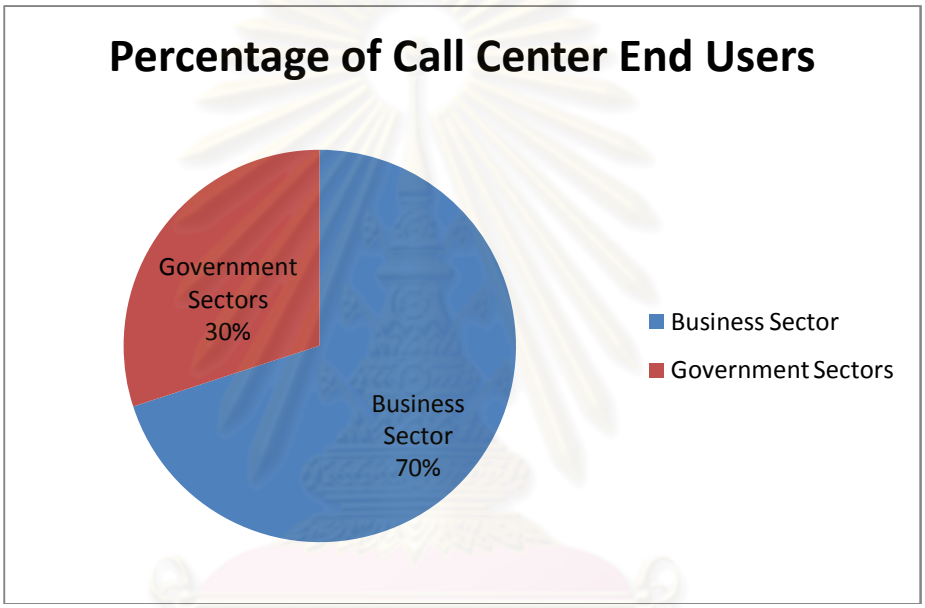


Figure 52 Percentage of Call Center Customers (According to Kasikorn Research Center)

According to Frost & Sullivan, the telecommunication and banking & finance are the top 2 industry verticals that use call centers, and the government and travel verticals ranked third and fourth, respectively. The growth should come from the telecommunication and banking & finance sector due to the expansions in the verticals in attempts to increase the service quality and database management efficiency. (Frost & Sullivan, 2008)

5.4.3 Outsourcing Trend in the Call Center Business

According to a key player interview of a call center software company, about 40% of the total end user companies in Thailand choose to do in-house call centers by buying only the infrastructures and software service, while 60% choose to outsource the

whole call center operations. (Figure 53) From the interviews, the large to medium size companies tend to operate in-house call centers and only buy the infrastructures and solution software due to the information confidentiality issues. As for the small to medium size companies, they tend to outsource the call center operations due to the cost-effectiveness of operations and lower investment on the call center function. There are increasing trend in the outsourcing of call center due to the initiatives of the small to medium size companies to decrease the cost and increase customer satisfaction.

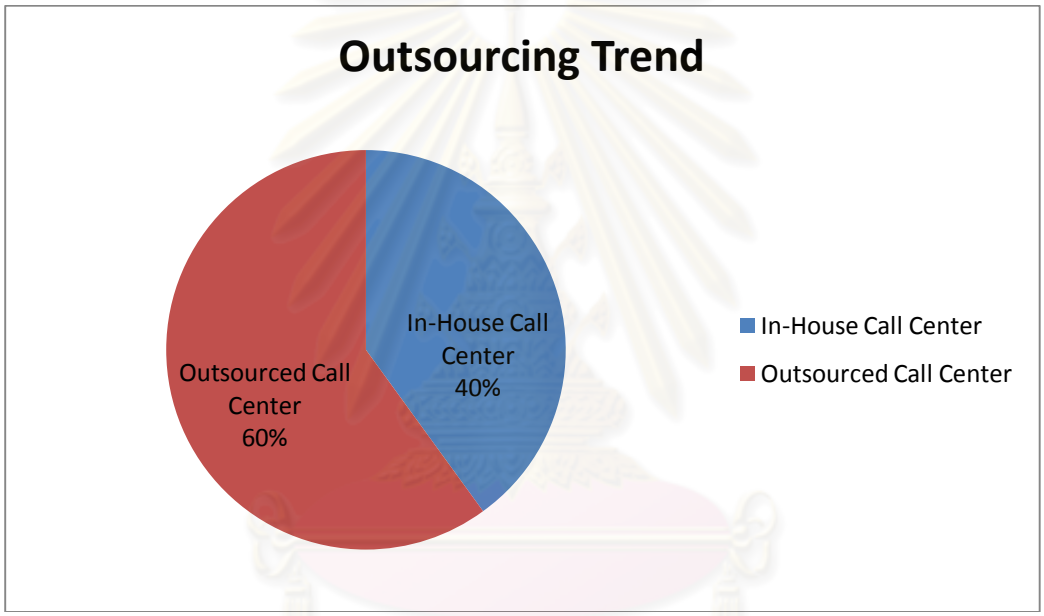


Figure 53 Outsourcing Trend in the Call Center Business (According to a Key Player Interview)

5.4.4 The Competition in the Call Center Business

According to Kasikorn Research Center, the competition in the call center business in Thailand is becoming more intense than in the past due to the increasing number of service providers and increasing number of international companies doing business in Thailand. The intense competition leads to the decrease in price for the call center solutions and services. (Kasikorn Research Center, 2009)

5.5 Market Study of Call Center Business

5.5.1 Overview of Interview on Call Center Software Developers

The in-depth interview on ten call center software developers were carried out. Some interviews were carried out by phone interviews, and some focus group interviews with the key players in the call center business were carried out. The interview focused on the implementation of the speech technologies in call center business. Some of the companies interviewed are the key players in the call center software business.

5.5.2 Companies Profile

From the software developers interviewed, there are three main types of developer:

1. The first type is the software houses that only develop the applications / solutions.
2. The second type of developer develops and also provides outsourced call center services (providing the human agent and manage the call center function of the end users).
3. The last type is the end user companies themselves that have enough IT competencies to develop the call center software and also manage the call center function by themselves.

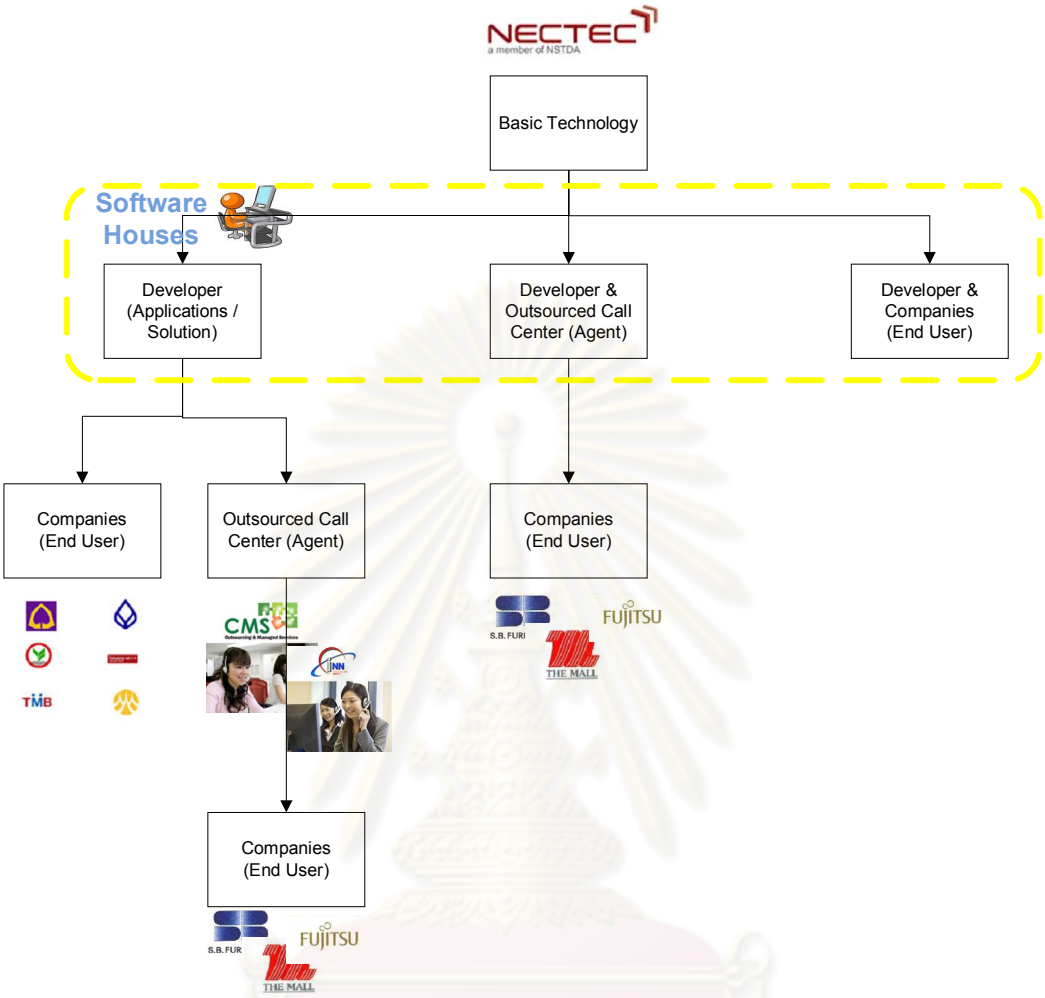


Figure 54 Value Chain of Speech Technology in Call Center Business

According to Figure 55, from the three types of developer, six interviews were conducted on the software houses that only develop and sell the applications / solutions to either the end user companies themselves or the outsourced call center service providers that will provide call center function services to the end user companies. One interview was conducted on the developers that also provide outsourced services. And three interviews were conducted on the end user companies that manage their own call center function and also develop the software by themselves

In Thailand business environment, the most abundant group of potential customers is the software houses that are specialized in developing call center

solutions. This study will mostly focus on the first and second group of customers (the software houses and the outsourced parties that develop their own software).

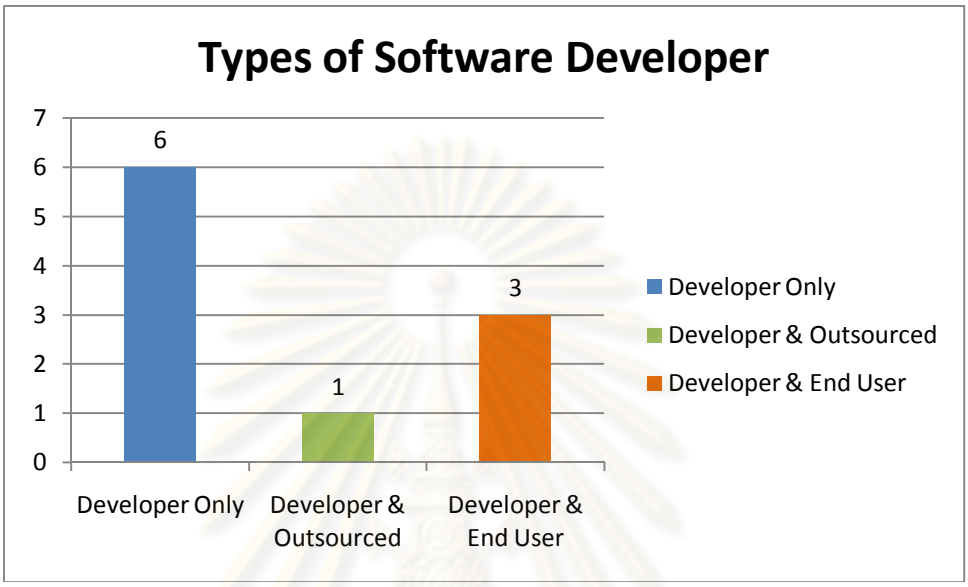


Figure 55 Types of Software Developer

5.5.3 Interest in Speech Technologies

According to Figure 56, from the interview, 9 out of 10 software developers were interested in speech technologies and have the plan to implement speech technologies with their current applications. There are two companies that also develop the speech technology software by themselves (one company collaborate with Nuance Communication and one is Thai company). This represents a considerably high interest rate in speech technologies for the call center software business. The outlook of speech technologies in the call center software business is promising. The only one company that does not have interest in implementing speech technologies is an end user company that develops the software by itself. The reason that they do not want to implement speech technologies is that the speech technologies do not suit the nature of their company's call center function. Since this company serves only the enterprise customers, all inquiries are complex inquiries that require engineers to respond. General

tasks that can be completed by speech technologies are not a part of this company's call center tasks.

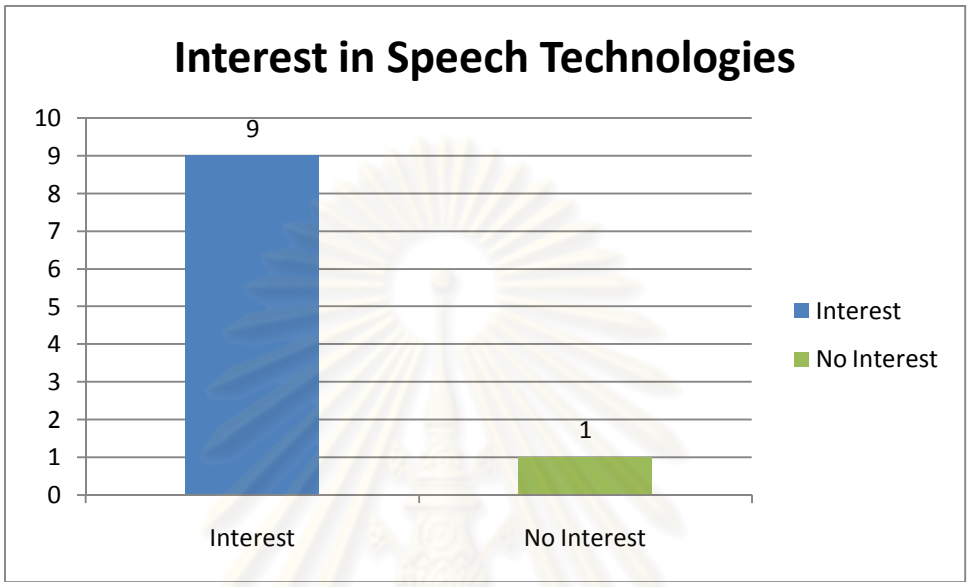


Figure 56 Interest in Speech Technologies

5.5.4 Types of End User

According to Figure 57, from the interview, the companies may serve public or private sectors, or both. There are a greater number of private end user companies comparing to the public sectors. This supports the finding of Kasikorn Research Center that the business sectors made up 70% of all call center end users, and the public sectors made up the leftover 30% (suggested in the 'Status of Call Center Business in Thailand' section).

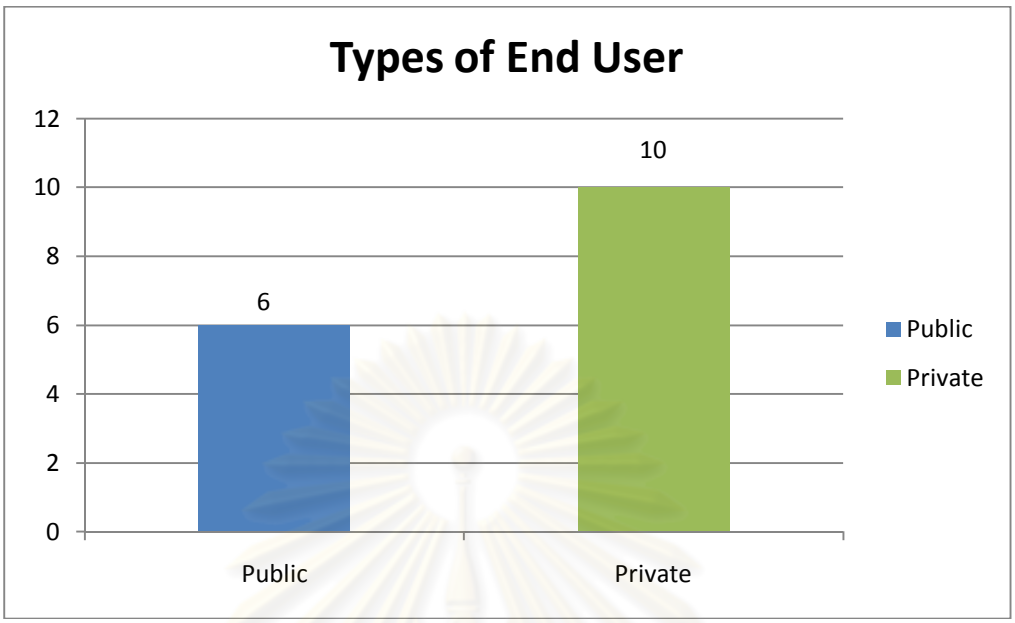


Figure 57 Types of End User

According to Figure 58, there are more large size private companies comparing to medium size. There are no small size companies. This number supports the information from one key call center software company that suggested that the large to medium size companies tend to operate in-house call centers and only buy the infrastructures and solution software due to the information confidentiality issues. As for the small to medium size companies, they tend to outsource the call center operations due to the cost-effectiveness of operations and lower investment on the call center function (Suggested in the 'Status of Call Center Business in Thailand' section). Therefore, most of the small size companies would not buy only the software and operate the call center function by themselves, but rather would outsource the whole call center function.

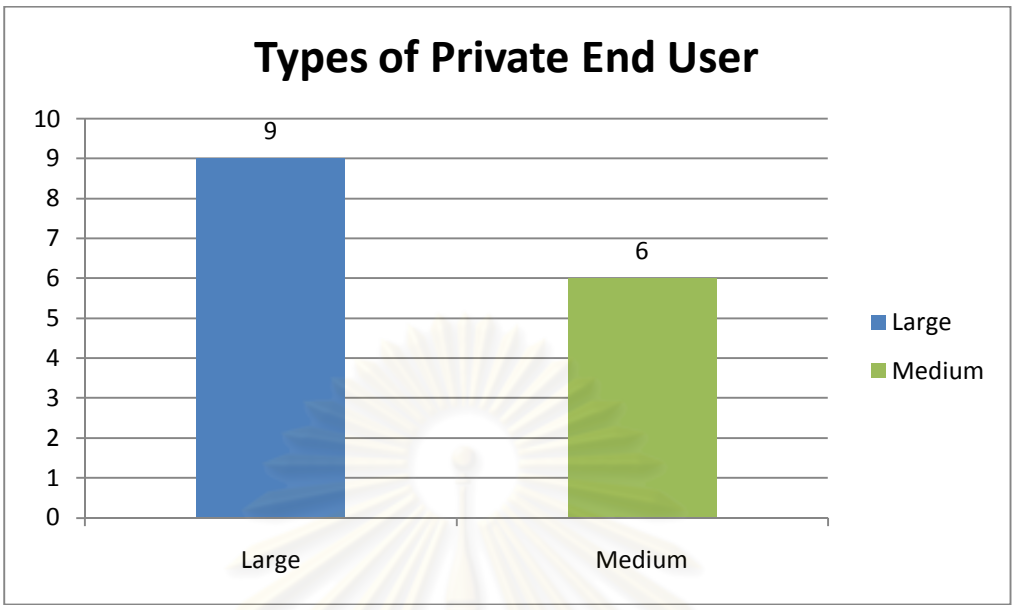


Figure 58 Types of Private End User

According to Figure 59, the end user industry verticals that are the top three sectors are: telecommunications, financial, and the travel sector. This supports the finding by Frost & Sullivan that the telecommunication and banking & finance are the top 2 industry verticals that use call centers, and the government and travel verticals ranked third and fourth, respectively (suggested in the 'Status of Call Center Business in Thailand' section).

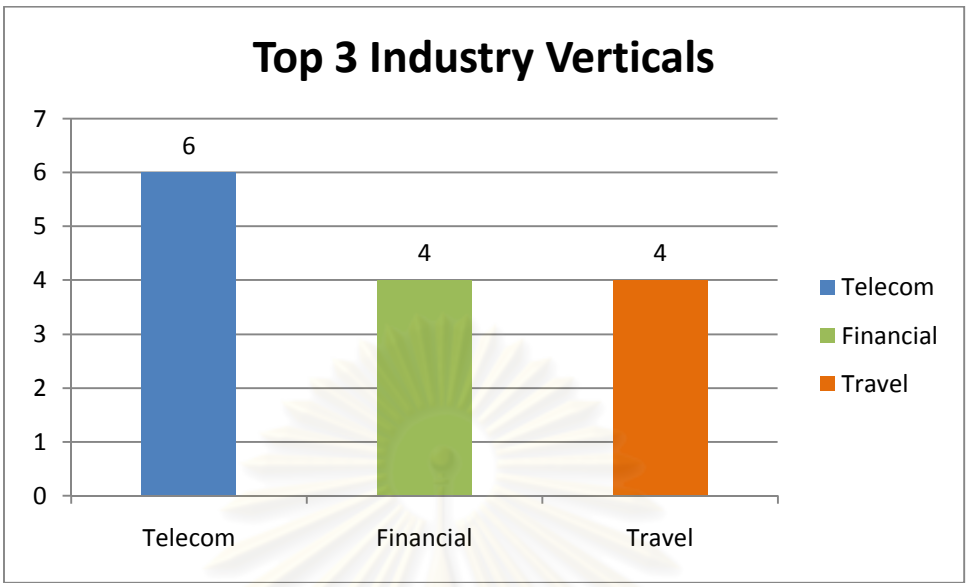


Figure 59 Top Three Industry Verticals

5.5.5 Technologies that Companies are Interested

The companies are asked about which type among the three types of speech technologies that they are interested in:

1. Text-to-Speech (TTS)
2. Automatic Speech Recognition (ASR)
3. Automatic Speaker Verification (ASV)

According to Figure 60, most interested companies are interested in TTS and ASR technology comparably with 43% and 50% interest rate respectively. However, there are 7% are interested in ASV technology as well.

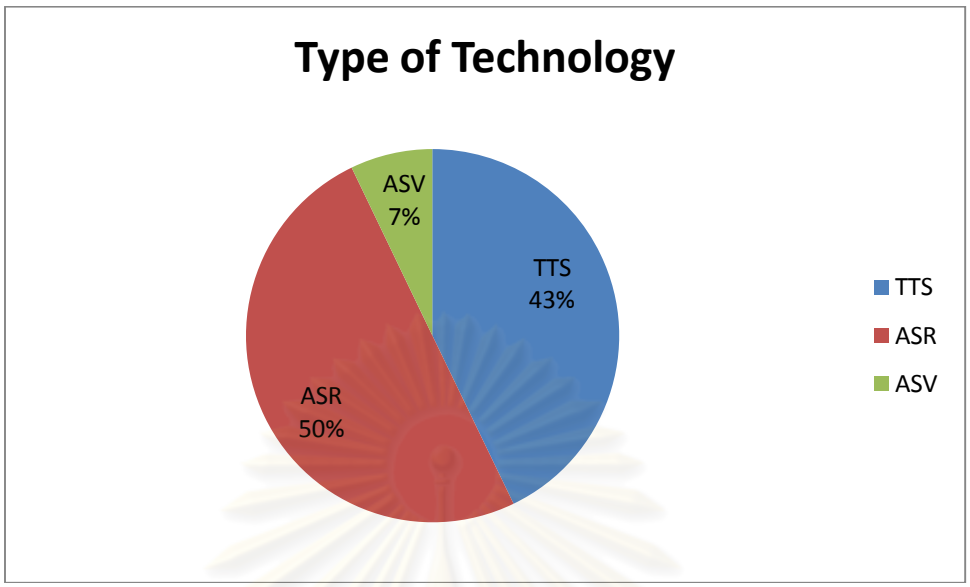


Figure 60 Type of Technology

5.5.6 Suitable Type of End User

According to the interview, all the developer companies suggested that the speech application is suitable for call centers in all industries. However, some suggested that the customers must be large enough to afford the extra fees or an increase in price of the call center solution.

5.5.7 Details of Services

The services that these developers plan to implement speech technologies are the support to the traditional IVR system. General tasks can be taken over by speech technology system to reduce the need of human operators for simple and repetitive tasks. The suggested plan of service to implement speech technologies in call centers are as follows:

Text-to-Speech (TTS)

TTS system is suggested to be used as a replacement to the pre-recording voice that can provide repetitive, dynamic, and frequently changing information. Some types of information that will be suitable for using TTS are:

- General information about the company

- Frequently asked questions
- Movie, train, airline schedules
- Customer's information such as customer's names, addresses, account balance

The information can be accessed through the database system and read to the callers without the need of human operators. This will benefit in terms of removing all general and repetitive tasks for the operators, which will ultimately increase the efficiency by increasing the contact channels during rush time. Also, the system will be able to respond to the callers 24/7, and could potentially work during the night time as well.

Automatic Speech Recognition (ASR)

ASR system is suggested to be used as an alternative to the traditional touchtone system (keypads). During the initial phrase of implementation, the choices between inputting by voice (ASR system) or by keypads (touchtone system) should be given since ASR system has rarely been implemented in Thailand. The callers should be able to get used to the ASR system before it is fully implemented by giving them options because they might not be accustomed to it in the first place and just turn away from it. If the callers became more accustomed to it and ASR becomes the standard implementation, the full implementation can be carried out.

The plans to implement ASR system are to take input by voice such as menu bypassing, voice search, call routing, and input by voice.

The bypassing of the menu is one of the most suggested services that are suitable to implement ASR system. The long list of menu can be avoided by using voice command, which will increase the customer satisfaction because they do not have to listen and remember which number to dial.

Voice search is also a practical application of ASR to search for a specific data such as voice search for music downloading.

The call routing to the specific function or department is another practical service. It allows the callers to just say the contact person's name and department and the call will be automatically routed. This removes the necessity of human operators for the simple and repetitive tasks of call routing.

Input by voice can be implemented to input the information that cannot be inputted by keypads such as caller's name or address. Also, if a lot of information is required, ASR is a suitable means of inputting the information.

In summary, the types of information that will be suitable for using ASR are:

- Long list of menu
- Search of information or specific items
- Data that cannot be inputted by keypads (names, addresses)
- Data that requires a lot of input

By implementing ASR system will eventually increase the customer satisfaction due to the increasing convenience, efficiency, and wider range of capabilities.

5.5.8 Suggested Licensing Agreement with NECTEC

From the interview, the developer companies were asked about how they preferred the form of licensing agreement between paying one-time upfront fee and the revenue share. The one-time upfront fee has the advantage for the customers that they do not have to continuously pay for NECTEC; they only need to pay only if they needed extra services. However, the disadvantage of upfront agreement is the large amount of investment required.

As for the revenue share, the advantage for the customers is that they do not need to pay a large amount of investment fee. The companies that selected for the revenue sharing suggested that they could lower the risk of implementation, and if it does not work out, they lose less than paying a large upfront. Also, they suggested that by using the revenue sharing scheme, NECTEC would be their partner rather than their supplier. They will gain in the way that NECTEC would help them carry out the

implementation. However, the disadvantage is that they would have to be committed with NECTEC for a long time, and it could be less flexible than paying one-time fee.

According to Figure 61, from the interview, there are seven responses for this topic, and they are approximately equally shared between the two schemes. Most of the larger companies responded for the upfront since they do not want to continuously pay the share. Also, they are large enough in capital and capability that they are certain to make profit from the investment. However, two out of three of the companies that responded for revenue share are small companies. Their reason for choosing revenue share is that they do not want to be fully exposed to the risk to invest in a large amount of money. Also, they prefer NECTEC to be their partner and share the risk of implementation as well as providing them with extra helps and services. Currently, NECTEC uses a combination of upfront and revenue sharing scheme.

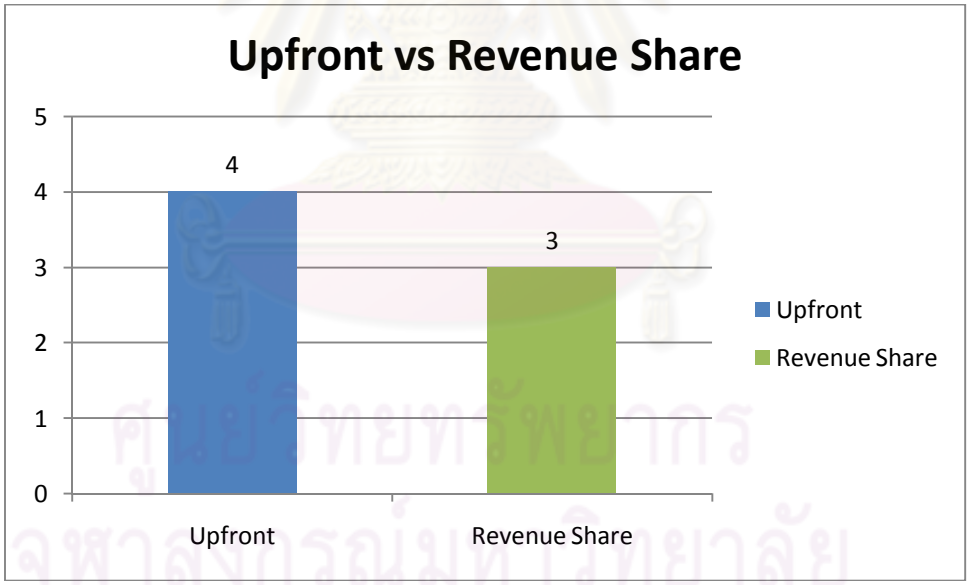


Figure 61 Upfront vs Revenue Share

5.5.9 Suitable Marketing Strategy

The developer companies were asked about their views of the suitable marketing strategy that they will use with their customers. The first strategy is to include speech technologies to the standard solution package and increase the price for the

whole package. The second strategy is to provide speech technologies as an option to the standard solutions. The third strategy is to provide speech technologies as a value-added gimmick for the standard solutions, free of cost. According to Figure 62, most of the companies suggested that they either would provide speech technologies as an option or free of cost. Most companies see the speech technologies as an innovative value-added function that might increase the customer satisfaction.

Figure 62 Suitable Marketing Strategy

5.6 Assessment of NECTEC's Capabilities Comparing with Call Center's Requirements

The interview asked about the capabilities of TTS and ASR software. The developer companies were asked if they have experience testing NECTEC's TTS software (VAJA 6.0) and ASR software (iSpeech). If they do not have experience with NECTEC's speech technologies, they will be asked about their opinions toward overall speech technologies in Thailand that they have encountered.

5.6.1 TTS Capabilities

The interview asked about three main aspects of capability: speech quality, naturalness, and speech intelligibility. The total interview was based on 8 companies, 6

of them have experience with NECTEC's TTS system (VAJA 6.0), and 2 of them have experience on other TTS software.

Speech Quality of TTS

According to Figure 63, looking at the overall TTS technology (NECTEC's software and other software), most companies think that the current speech quality of TTS technology in Thailand is in the acceptable range (7 out of 8 opinions), and one opinion thinks that it is not yet acceptable (1 out of 8 opinions).

While for NECTEC's TTS software (VAJA 6.0) the developer companies that have tested it all think that the speech quality is acceptable (6 out of 6 opinions).

This suggests that the perception toward speech quality of TTS system as a whole is considered good, especially with the NECTEC's TTS technology.

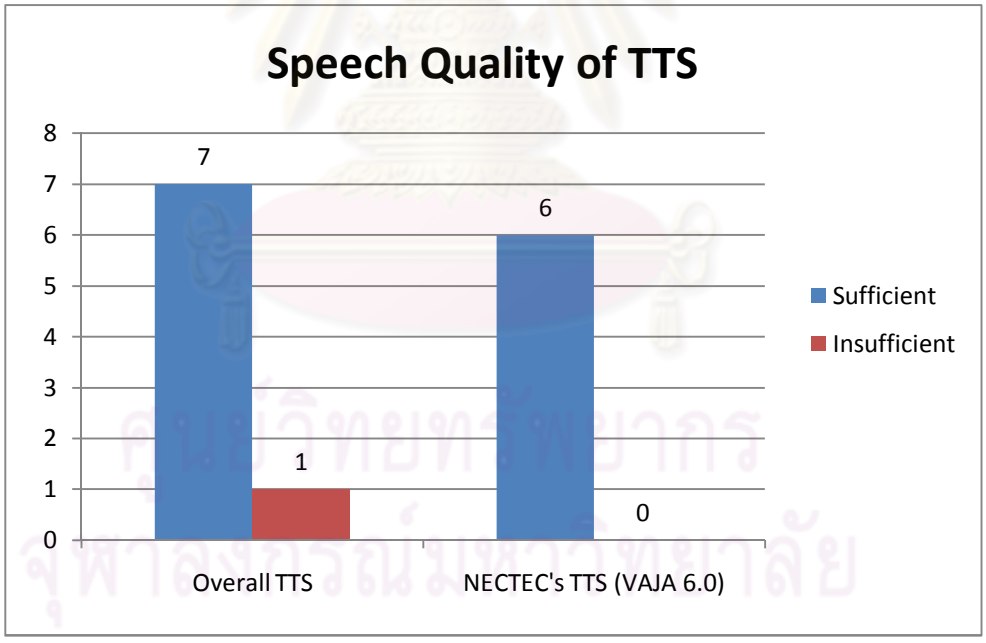


Figure 63 Speech Quality of TTS

Naturalness of TTS

According to Figure 64 looking at the overall TTS technology (NECTEC's software and other software), there are equal number of companies that think the

naturalness is in the acceptable range and that think it is not acceptable (4 out of 8 opinions).

While for NECTEC's TTS software (VAJA 6.0) there are 4 developer companies that have tested it and think that the naturalness is already sufficient (4 out of 6 opinions), and 2 companies think that the naturalness must be improved (2 out of 6 opinions).

This suggests that the perception toward naturalness of TTS system as a whole is considered average, but the score for NECTEC's is better. This might suggest the perception from the customers of NECTEC's naturalness might be better than other TTS system in Thailand.

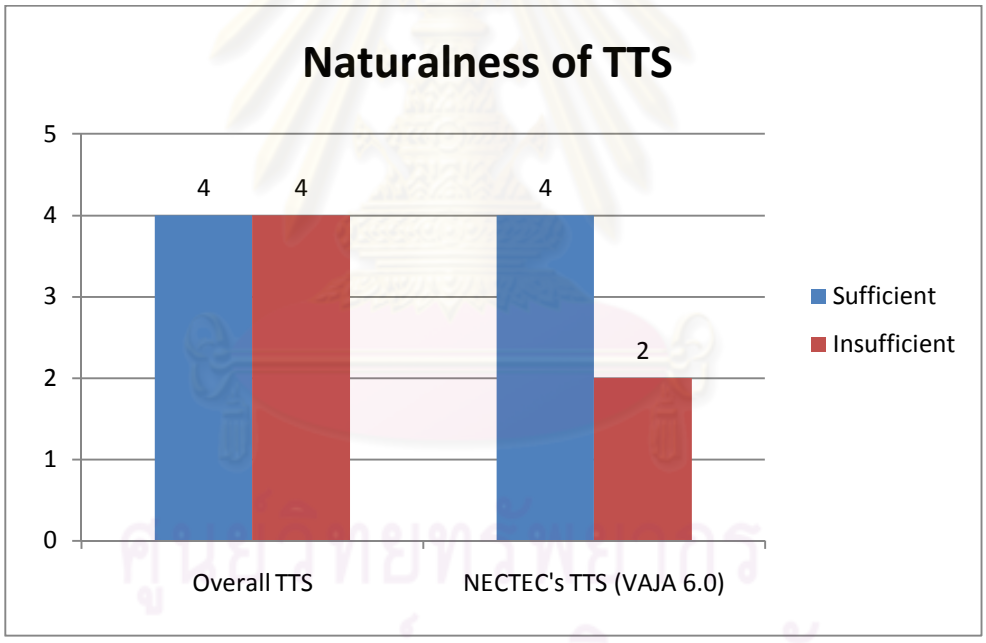


Figure 64 Naturalness of TTS

Speech Intelligibility of TTS

According to Figure 65, looking at the overall TTS technology (NECTEC's software and other software), there 6 companies that think the TTS system can read the text correctly (6 out of 8 opinions), 2 companies that think the TTS system still cannot read the text correctly enough (2 out of 8 opinions).

While for NECTEC's TTS software (VAJA 6.0) there are 5 developer companies that have tested it and think it can read the text correctly (5 out of 6 opinions), and one company thinks that the naturalness must be improved (1 out of 6 opinions).

This suggests that the perception toward speech intelligibility of TTS system as a whole is considered average, but the score for NECTEC's is a little bit better. This might suggest that NECTEC's intelligibility might be better than other TTS system in Thailand.

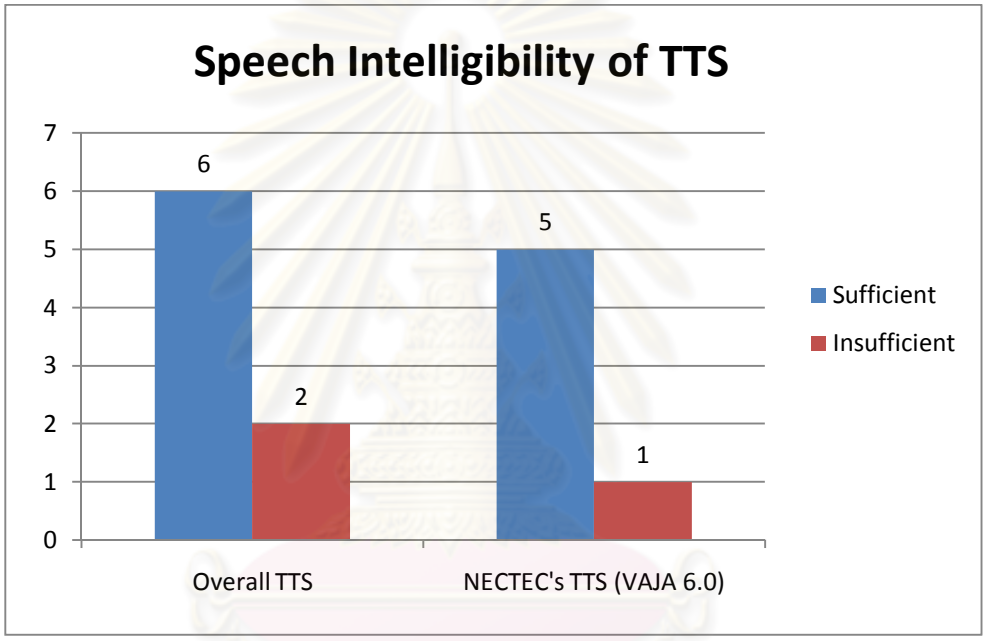


Figure 65 Speech Intelligibility of TTS

Conclusion on TTS Capabilities

When comparing between the opinions toward NECTEC and toward overall TTS system in Thailand, NECTEC seems to score better. This might suggest that NECTEC's TTS system might perform better since the ones who already tested mostly suggested that the system is sufficient in capabilities.

In terms of speech quality, NECTEC already perform well. Some concerns might be about the naturalness and speech intelligibility, and some efforts might be increased in order to improve these two areas. However, most opinions toward NECTEC's TTS system suggest that its capabilities are sufficient enough to be commercialized.

5.6.2 ASR Capabilities

The interview asked about three main aspects of capability: accuracy, required domain, and required working environment. The total interview was based on 8 companies, 3 of them have experience with NECTEC's ASR system (iSpeech), 4 of them have experience on other ASR software, and 1 of them never have any experience on ASR system in Thailand. The companies that do not have experience with NECTEC's ASR technology or never have any experience on ASR technology in Thailand will be asked about their opinions toward overall ASR technology in Thailand that they have encountered.

Accuracy of ASR

According to Figure 66, looking at the overall ASR technology (NECTEC's software and other software), most companies think that the current accuracy of ASR technology in Thailand is not in the acceptable range (5 out of 8 opinions), and 3 opinion thinks that it is acceptable (3 out of 8 opinions).

While for NECTEC's TTS software (VAJA 6.0) the developer companies that have tested it all think that the speech quality is not yet acceptable (3 out of 3 opinions).

This suggests that the perception toward accuracy of ASR system as a whole is considered not good, especially with the NECTEC's ASR technology. Most company requires high accuracy from ASR system, and the current accuracy level of the system in Thailand (including NECTEC's) is not up to standard.

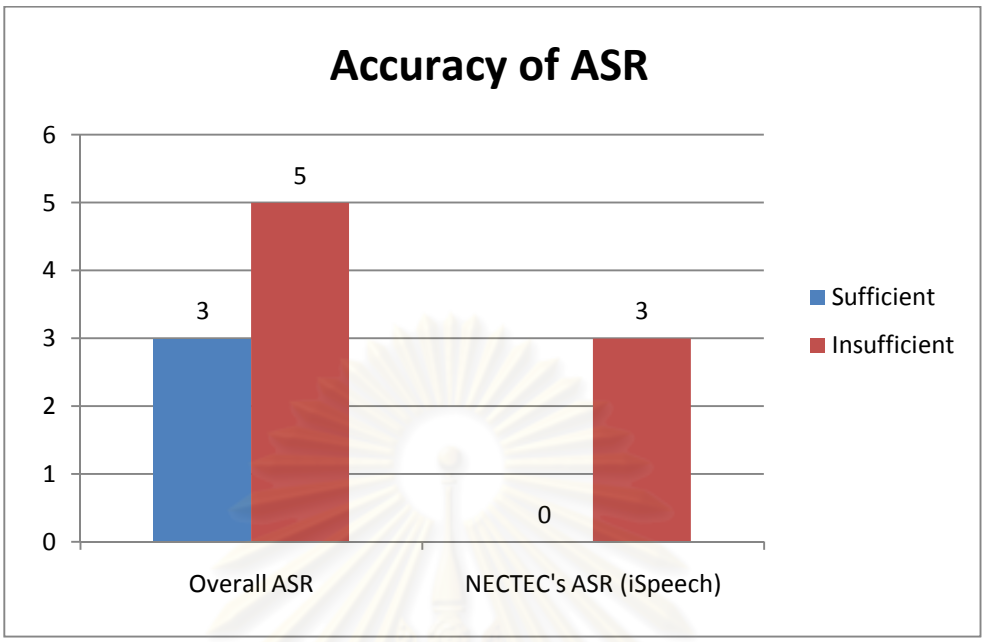


Figure 66 Accuracy of ASR

Required Domain for ASR

About the required domain level for ASR system, the companies were asked about their desired level of domain input. The domain levels were categorized into three main levels:

- Isolated Words- For example, voice command.
- Sentences (Continuous Speech with Limited Grammar) - Usually uses in spoken dialog system that can recognize a phrase or sentence that is within the selected scope.
- Continuous Conversation (Continuous Speech with Unlimited Grammar) - Any possible conversation likes human's conversation such as dictation and transcription.

Currently, NECTEC's ASR technology is limited to the isolated word level. According to Figure 67, most of the developer companies only require the isolated word level of input (4 out of 8 opinions). They gave their reasons that they think only isolated words are enough to input the demand. However, some of them suggested that it would

be a plus if sentences can be used. With a close score to the isolated words, some companies require that the sentences level is achieved (3 out of 8 opinions). The reason is that they think isolated words are not enough to impress and create customer satisfaction, but they think continuous conversation is unnecessary. The last level is the continuous conversation, which is not required by most developers. There is only one company that requires continuous conversation (1 out of 8 opinions).

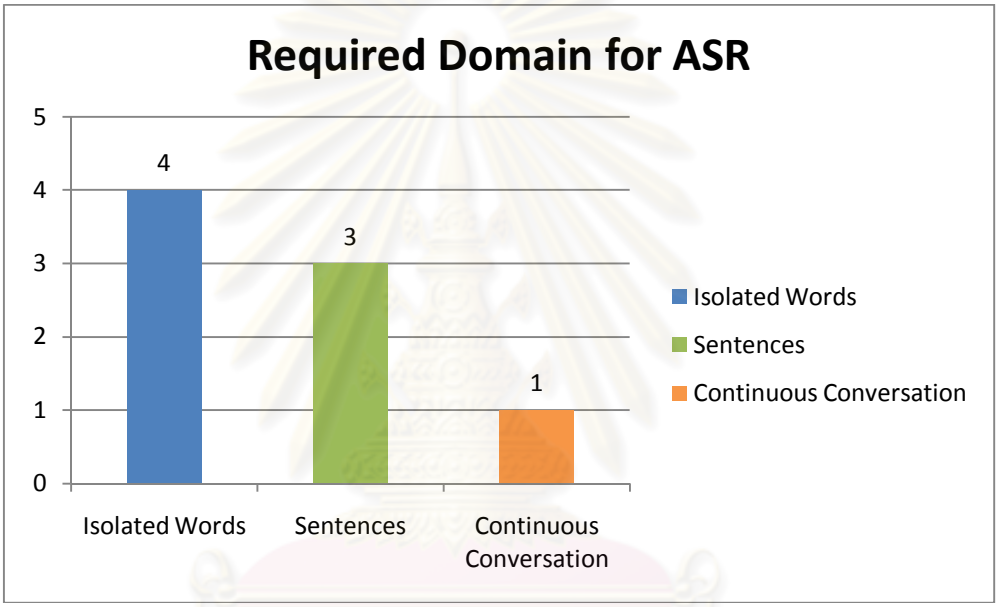


Figure 67 Required Domain for ASR

Required Working Environment for ASR

The companies were asked about their desired environment that ASR system will be able to operate. The environments were categorized into three main environments:

- Quiet Room
- Office
- Outdoor

Currently, NECTEC's ASR technology is limited to the office environment. According to Figure 68, most of the developer companies require the ASR system to be able to operate at any environment; therefore, most of them selected the outdoor

environment (6 out of 8 opinions). They gave their reasons that they think the users should be able to call from any place at anytime. Some of the companies said that the office environment would be enough since most of their users will be in the office environment (2 out of 8 opinions). The last level is the quiet room, which all the companies think it would not be likely that the users will call from a quiet room (0 out of 8 opinions).

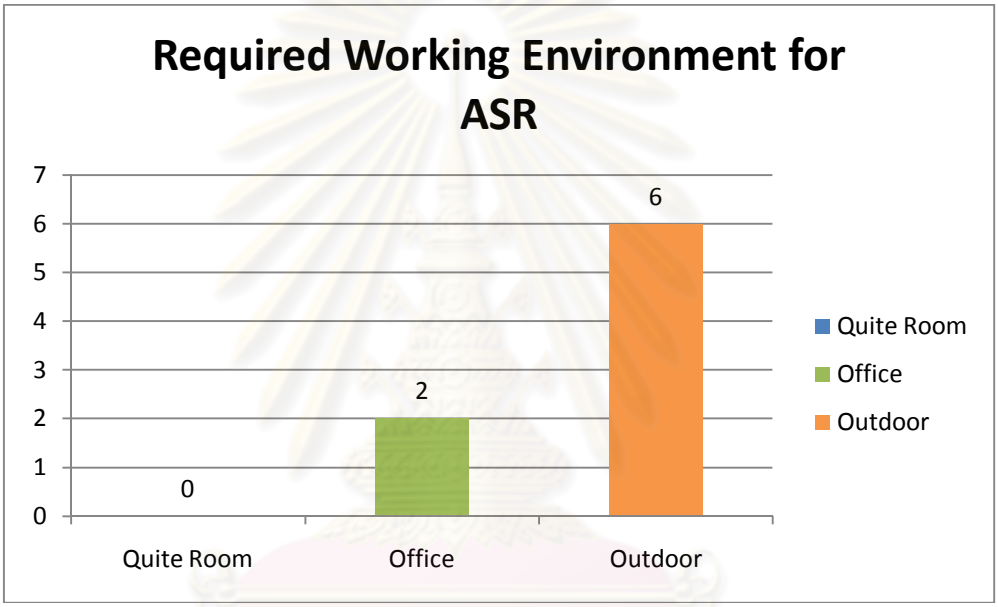


Figure 68 Required Working Environment for ASR

Conclusion on ASR Capabilities

From the interview, it could be concluded that NECTEC's ASR system (iSpeech) might not be suitable for commercialization at the present. Looking at the accuracy issue, the developer companies are not confident about the accuracy and most think that the current accuracy level is not enough. For the domain aspect, even though most companies only require the isolated words, it would be better to improve to the sentence level. As for the working environment, most companies require that the ASR is able to be used at any environment. Currently, it could only be used in the office environment, which is not enough for most customers.

Therefore, NECTEC should first focus on improving the accuracy issue and the working environment to be able to operate outdoor. Then if both of the issues suggested are improved, efforts can then be moved to improve the domain level to the sentence level.

5.6.3 Conclusion for NECTEC's Speech Technologies' Capabilities

From the interview about the customers' views toward NECTEC's and overall speech technologies in Thailand, NECTEC's TTS software (VAJA 6.0) has the potential to be commercialized at present. The focus might be put on improving the naturalness and speech intelligibility to ultimately increase the customer satisfaction. As for ASR, most customers still doubt about the accuracy issue, both in overall system and NECTEC's system (iSpeech). At present, NECTEC should improve the accuracy for the system with the current domain level of isolated words. Also, for the working environment that ASR system can be used with, NECTEC should improve from office level to outdoor level. Since most customers require that they would want their users to be able to use at any place anytime, it is one important criterion that they would require. Then if both of the issues suggested are improved, efforts can then be moved to improve the domain level to the sentence level in the future.

5.7 Study of End Users

5.7.1 Overview of Interview on End Users of Call Center Software

The interviews on four call center software end users were carried out. The end users are companies or sectors that purchase the call center software and services from the software houses and implement their own call center function with the software. The interviews were carried out by phone interviews with the managers of the call center function or related position. The interview asked if the end users have the demand in speech technologies for their call center function.

5.7.2 End User Companies Profile

From the end users interview, there were one public enterprise and four private enterprises. Two of the private enterprises were large size private enterprise and two were small size private enterprise. The industries that the sectors operate in are:

1. Public telecommunications business
2. Large private telecommunications business
3. Large private retailer business
4. Small private IT system business
5. Small private IT system business

5.7.3 Awareness and Interest in Speech Technologies

According to Figure 69, out of five companies interviewed, four of the companies have some knowledge about speech technologies, and one have never heard about speech technologies.

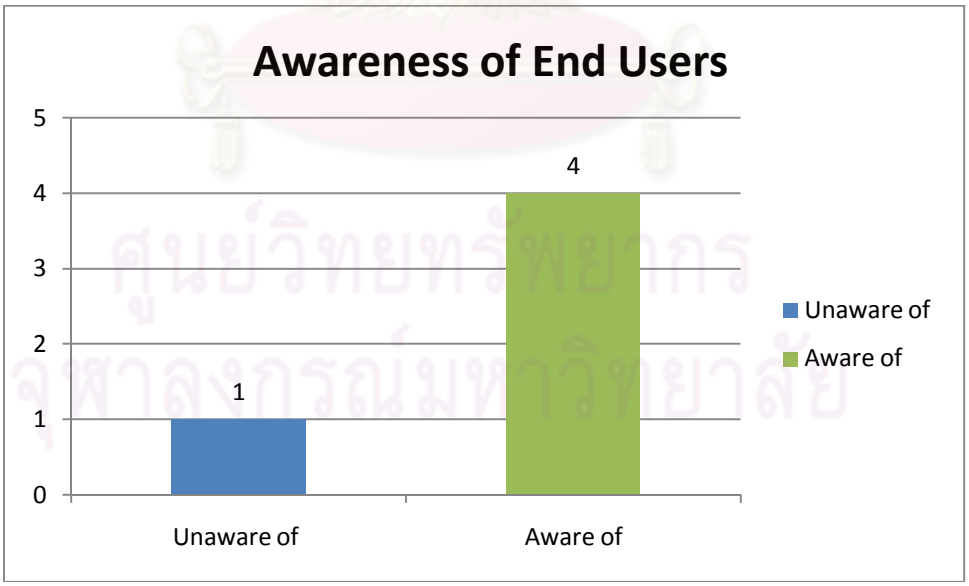


Figure 69 Awareness of End Users

The companies were then asked about the interest of using speech technologies for their call center function. For the company that never heard about speech technologies, a brief explanation was given about the technologies, and they were asked if they see the potential for speech technologies to be used with call center function (without considering the cost added in the first place). According to Figure 72, all five companies have interest in speech technologies.

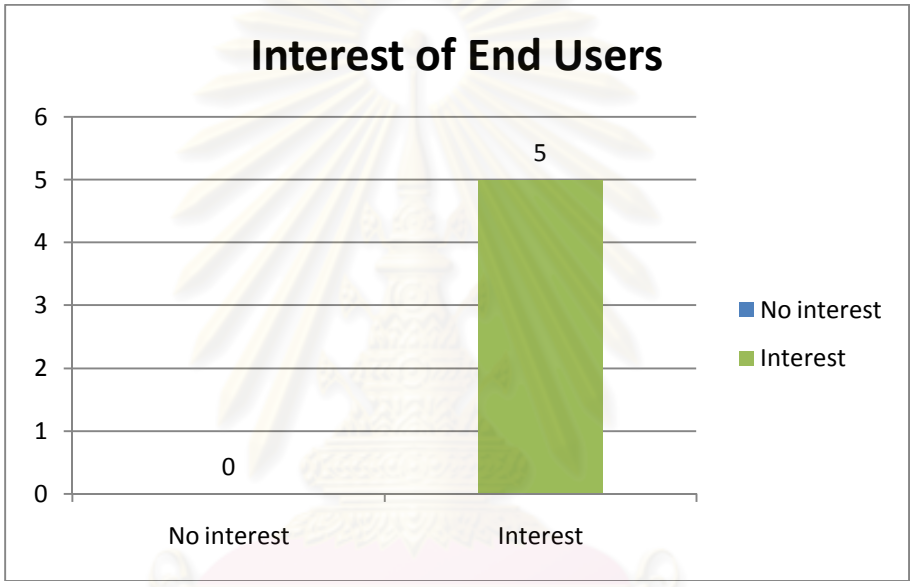


Figure 70 Interest of End Users

The companies gave their reasons of why they want to employ speech technologies with their call center functions as follows:

- Could potentially add value and differentiate their services
- Can increase the efficiency of the process
- Create innovative image for their corporate
- Increase customer satisfaction
- Can be used with some specific groups of users such as handicapped and old people

5.7.4 Willingness to Pay for Investment

The companies were asked if there are extra costs for speech technologies added to the existing product / service will they be willing to pay for the extra costs, and at which level. Four companies responded that they are willing to pay for the extra costs if they see that the benefits are greater than the costs, they are willing to pay for it. There are no fixed limits of the extra costs they are willing to pay, but they think that the benefits in terms of creating innovative image might overcome the costs. One of the four companies responded that the extra costs should be paid by installments in smaller amounts rather than one large investment. However, a small company responded that they would like to see some more research or technology pacesetters to confirm that the implementation will be successful.

5.8 Approximate Market Worth of Call Center Software Business (2009)

According to Kasikorn Research Center, the total market worth for call center business in Thailand in the year 2008 was approximately 3,000-3,500 million THB (approximately US$90-110 million). (Kasikorn Research Center, 2009) This number includes the revenue from all sectors of the call center business:

1. Outsourced call center service- Provide the whole solutions and infrastructure of call centers with the agent service for companies to outsource call center function.
2. Call center system and software service- Provide the solutions of call centers and infrastructures to companies to do in-house call center function.
3. Telemarketing- Outbound call centers with the telemarketing function.

In this part of the study, ***the approximation of the market worth of only the call center software business*** is carried out in order to anticipate the outlook of speech technologies in the call center software business.

The approximation is carried out based on the information from NECTEC about the existing call center software companies. It is possible that there might be few more

companies that are not in the list, but this source of information is considered reliable to some extent. From the information, there are approximately 28 existing call center software companies in Thailand.

5.8.1 Procedures on the Approximation of the Market Worth

1. The revenue of each company is gathered from Thailand's Department of Business Development website (www.dbd.go.th), which provides the financial statement including the balance statement and profit and loss statement. The latest information provided is the statement for 2009. Therefore, this approximation is the approximate market worth of 2009.
2. The revenue of the call center software from each company is calculated by taking only the proportion of the call center business to the total business (for example, only 40% of a company operates in the call center business, while another 60% is in other businesses; therefore, only 40% of the total revenue is taken into account). Also, if a company is also an outsourcing provider, the revenue from only software is taken into account (for example, only 20% of the total revenue came from the software, while another 80% came from the outsourcing service; therefore, only 20% of the total revenue is taken into account).
3. The companies are categorized into three sizes: small (S), medium (M), and large (L), based on their authorized share capital. The criterion is as follows:
    - Small (S) : Authorized share capital of less than 3 million THB (< 3 million THB)
    - Medium (M) : Authorized share capital of between more than or equal to 3 million THB to less than or equal to 10 million THB ($\geq$ 3 million THB to $\leq$10 million THB)

- Large (L) : Authorized share capital of more than 10 million THB (>10 million THB)

4. The averages of the revenue from each category (S, M, L) are calculated. Since there are some companies that the information cannot be gathered, only 20 companies are taken into the calculation of the averages.
5. The total frequency of each category (S, M, L) are counted (including companies that the information cannot be gathered as well)
6. The market worth is calculated by the following formula:

$$Market\ Worth = (\bar{X}_S \times N_S) + (\bar{X}_M \times N_M) + (\bar{X}_L \times N_L)$$

$$Where\ \bar{X}_S = Average\ Revenue\ of\ Small\ Companies$$
$$\bar{X}_M = Average\ Revenue\ of\ Medium\ Companies$$
$$\bar{X}_L = Average\ Revenue\ of\ Large\ Companies$$

$$And\ N_S = Total\ Frequency\ of\ Small\ Companies$$
$$N_M = Total\ Frequency\ of\ Medium\ Companies$$
$$N_L = Total\ Frequency\ of\ Large\ Companies$$

5.8.2 Conclusion on the Market Worth of the Call Center Software Market

From the approximation, the market worth of call center software business in Thailand (2009) is approximately 595 million THB (approximately US$ 19 million). The detailed calculation is provided in Appendix B. This number suggests an interesting portion of the total call center business in Thailand, which can be a good segment for speech technologies to enter the business.

5.9 Approximate Market Worth That NECTEC Could Potentially Gain (Call Center)

According to Opus Research, the global speech technology market worth had reached US$ 1 billion in 2006. From that, the portion of the speech technology used in call center business was approximately US$ 600 million. (Opus Research, cited in Business 2.0 Magazine, 2007)

According to Business Insights, the global market worth for call center technology was estimated to be US$ 4.5 billion in 2006. (Furness, cited in Business Insights) This makes speech technology in the call center business worth 13% out of the total call center technology worth (US$ 600 million to US$ 4.5 billion). This can be summarized in the following figure:

Figure 71 Global Speech Technology in Call Center Market Worth

For year 2009, the call center software market worth in Thailand is estimated from the previous section to be 595 million THB (approximately US$ 19 million). The year-on-year growth of the call center business in Thailand is approximately 18.4%. (Frost & Sullivan cited in The Nation Newspaper, 2010) From this, it can be estimated

that the market worth is 704 million THB for 2010, 834 million THB for 2011, and 987 million THB for 2012.

Assuming NECTEC will begin to seriously commercialize speech technology in 2011. Using to this percentage (13%) to forecast the market worth for speech technology in call center business in Thailand in the year 2011. It could be forecasted that the market worth of speech technology in call center business in 2011 is about 108 million THB. Assuming the market worth of the two technologies (TTS and ASR) are divided equally, 54 million THB are contributed to TTS and ASR technology each. Using the approximation of the market share of NECTEC (approximated by the experienced NECTEC's researchers), the share of NECTEC's TTS technology is approximately 30%, which is 16.2 million THB. And the share of NECTEC's ASR technology is approximately 20%, which is 10.8 million THB. **The approximated possible market worth that NECTEC could potentially gain in 2011 is about 27 million THB**. This is summarized in the following figure:

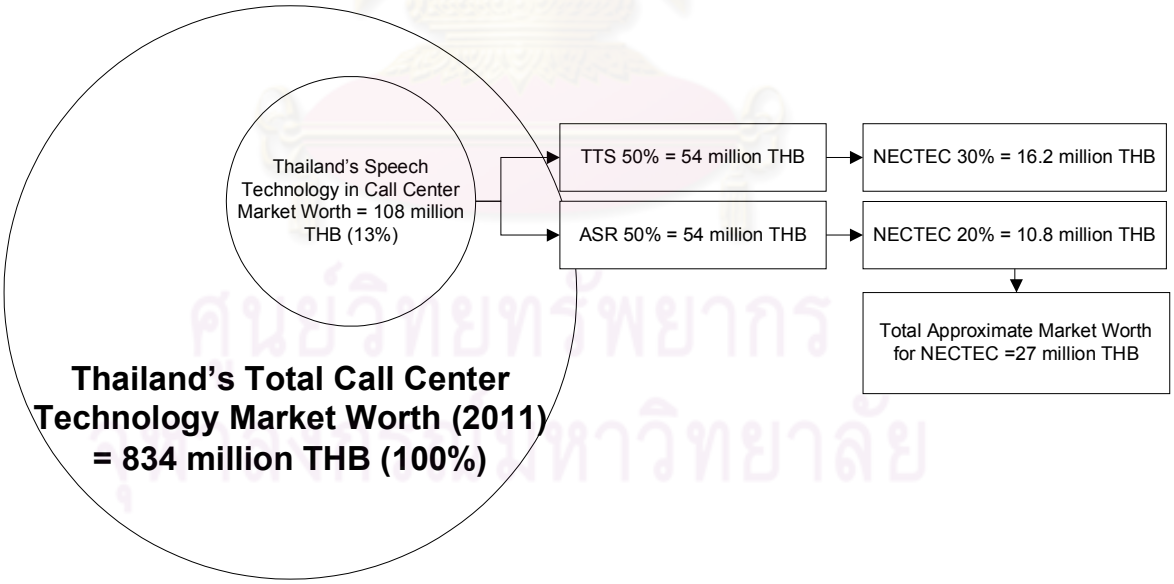


Figure 72 Thailand Speech Technology in Call Center Market Worth

However, if in the future NECTEC could improve the capabilities of the software and increase marketing and technology transfer, more share could potentially be

gained. The scenario of NECTEC gaining 50% share in both TTS and ASR technology in year 2012 can be summarized as follows with the potential market worth of 64 million THB.

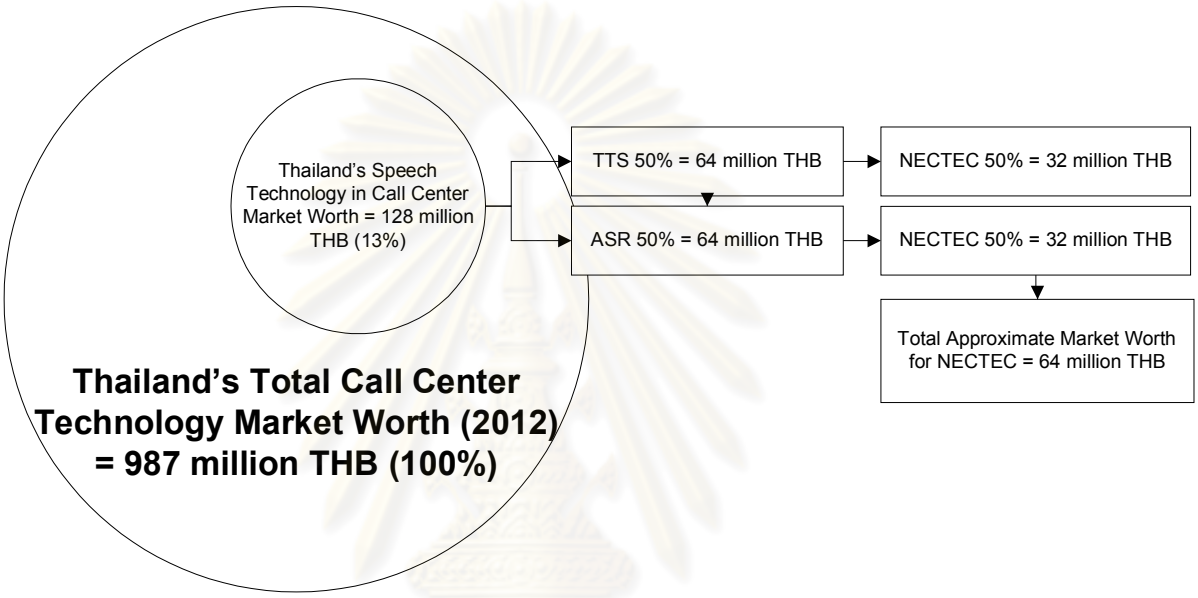


Figure 73 Thailand Speech Technology in Call Center Market Worth (Another Scenario)

5.10 Summary of Market Opportunities for Speech Technologies in Call Center Business

Nowadays, telephone is still the most heavily used of communication channel due to the convenience and the reasonable cost. Speech technologies can be implemented with IVR system in the call center function in order to:

- Improve customer satisfaction by convenience
- Bypass the menu
- Increase the range of data input and output
- Automate repetitive inquiries and transactions
- Increase efficiency and decrease operational costs
- Optimize human resource and their talents

The growth and the market worth of the call center business in Thailand attract many new companies to enter the call center business. As the competition increases, each business are trying to increase their competitive advantages by improving their performance, increase value-added functions, increase customer satisfactions, and build innovative image.

The call center business in Thailand can be segmented into three main service sectors: outsourced call center service, call center system and software service, and telemarketing service. (Kasikorn Research Center, 2009) In this study, the focus is mainly on the call center software service.

This study was conducted on three main types of call center software developer:

1. The first type is the software houses that only develop the applications / solutions.
2. The second type of developer develops and also provides outsourced call center services (providing the human agent and manage the call center function of the end users).
3. The last type is the end user companies themselves that have enough IT competencies to develop the call center software and also manage the call center function by themselves.

However, the most abundant group of potential customers for NECTEC is the software houses that are specialized in developing call center solutions. This study will mostly focus on the software houses and the outsourced parties that develop their own software.

From the interview, almost all of the call center software developers have interest in speech technologies and have the plan to implement speech technologies with their current applications.

A developer company may serve public or private sectors, or both. There were more private end user companies comparing to the public end users. According to the interview, most of the end users are large and medium size private companies. The end

user industry verticals that are the top three sectors are: telecommunications, financial, and the travel sector. According to the interview, all the developer companies suggested that the speech application is suitable for call centers in all industries. However, some suggested that the customers must be large enough to afford the extra fees or an increase in price of the call center solution.

The services that these developers plan to implement speech technologies are the support to the tradition IVR system. General tasks can be taken over by speech technology system and reduce the need of human operators for simple and repetitive tasks. TTS system is suggested to be used as a replacement to the pre-recording voice that can provide repetitive, dynamic, and frequently changing information. Some types of information that will be suitable for using TTS are:

- General information about the company
- Frequently asked questions
- Movie, train, airline schedules
- Customer's information such as customer's names, addresses, account balance

ASR system is suggested to be used as an alternative to the traditional touchtone system (keypads). During the initial phrase of implementation, the choices between inputting by voice (ASR system) or by keypads (touchtone system) should be given since ASR system has rarely been implemented in Thailand. The plans to implement ASR system are to take input by voice such as menu bypassing, voice search, call routing, and input by voice. The types of information that will be suitable for using ASR are:

- Long list of menu
- Search of information or specific items
- Data that cannot be inputted by keypads (names, addresses)
- Data that requires a lot of input

From the interview, the developer companies were asked about how they preferred the form of licensing agreement between paying one-time upfront fee and the revenue share. Most of the larger companies responded for the upfront since they do not want to continuously pay the share, and they are large enough in capital and capability that they are certain to make profit from the investment. The smaller companies chose the revenue share because they do not want to be fully exposed to the risk to invest in a large amount of money.

The developer companies were asked about their views of the suitable marketing strategy that they will use with their customers. Most of the companies suggested that they either would provide speech technologies as an option or free of cost. Most companies see the speech technologies as an innovative value-added function that might increase the customer satisfaction.

The interview asked about the capabilities of TTS and ASR software for overall speech technologies in Thailand and for NECTEC's software. For TTS, the interview asked about three main aspects of capability: speech quality, naturalness, and speech intelligibility. For ASR, the interview asked about three main aspects of capability: accuracy, required domain, and required working environment. From the interview, NECTEC's TTS software (VAJA 6.0) has the potential to be commercialized at present. The focus might be put on improving the naturalness and speech intelligibility to ultimately increase the customer satisfaction. As for ASR, most customers still doubt about the accuracy issue of NECTEC's system (iSpeech). At present, NECTEC should improve the accuracy for the system with the current domain level of isolated words. Also, for the working environment that ASR system can be used with, NECTEC should improve from office level to outdoor level. Then if both of the issues suggested are improved, efforts can then be moved to improve the domain level to the sentence level in the future.

The interview on the end users could potentially suggest that the end users have interest on speech technologies. However, it is noted that the limitations on the number of interviews can represent only a small part of the end users. Still, it could be

suggested that there are rooms for speech technologies in the call center business to grow. Most end users suggested that they are even willing to pay for extra costs if the investment will gain more in terms of benefits.

From all the study on the call center business, it can be suggested that the speech technologies have a promising future in the business, which is growing and still have rooms for a lot of improvements.

Chapter VI

Market Study of Future Potential Market

6.1 First Future Target: Accessibility Applications

6.1.1 Details for Accessibility Applications

The accessibility applications increase accessibility among people by using TTS and ASR technology. It is the second most interesting application with the highest interest rate along with the call center application. However, the accessibility applications are suitable to various types of business such as office & administrative software, manufacturing software, information management system, media & game, etc. Since there are many types of business that are suitable for the accessibility application, the customer base is large but hard to focus. In conclusion, accessibility application has very wide ranging solutions and can be used in a wide variety of businesses, but the drawback is that it would be hard to focus on a particular group of customer. In the future with more technology and knowledge transfer, the interest rate could be increased by giving the potential customers more information.

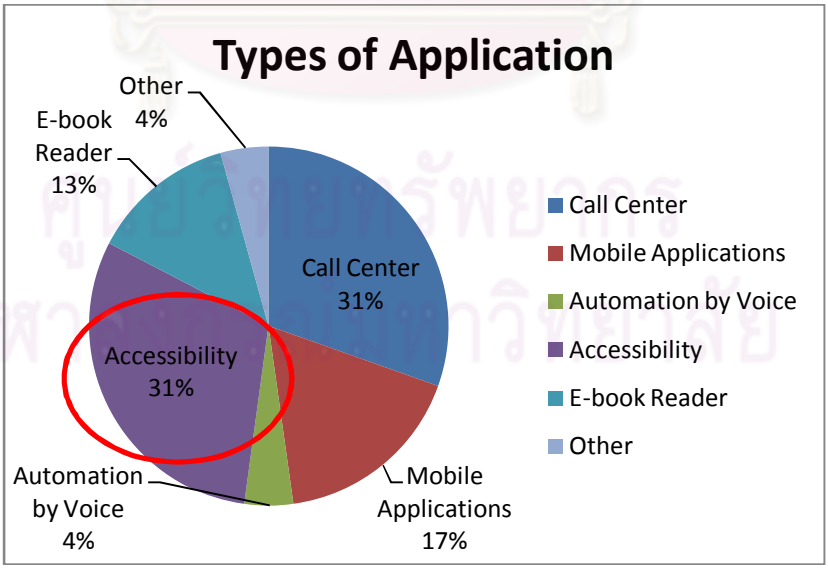


Figure 74 Types of Application

It can be seen that the interest in accessibility applications can be applied with various types of business as suggested in the figure below.

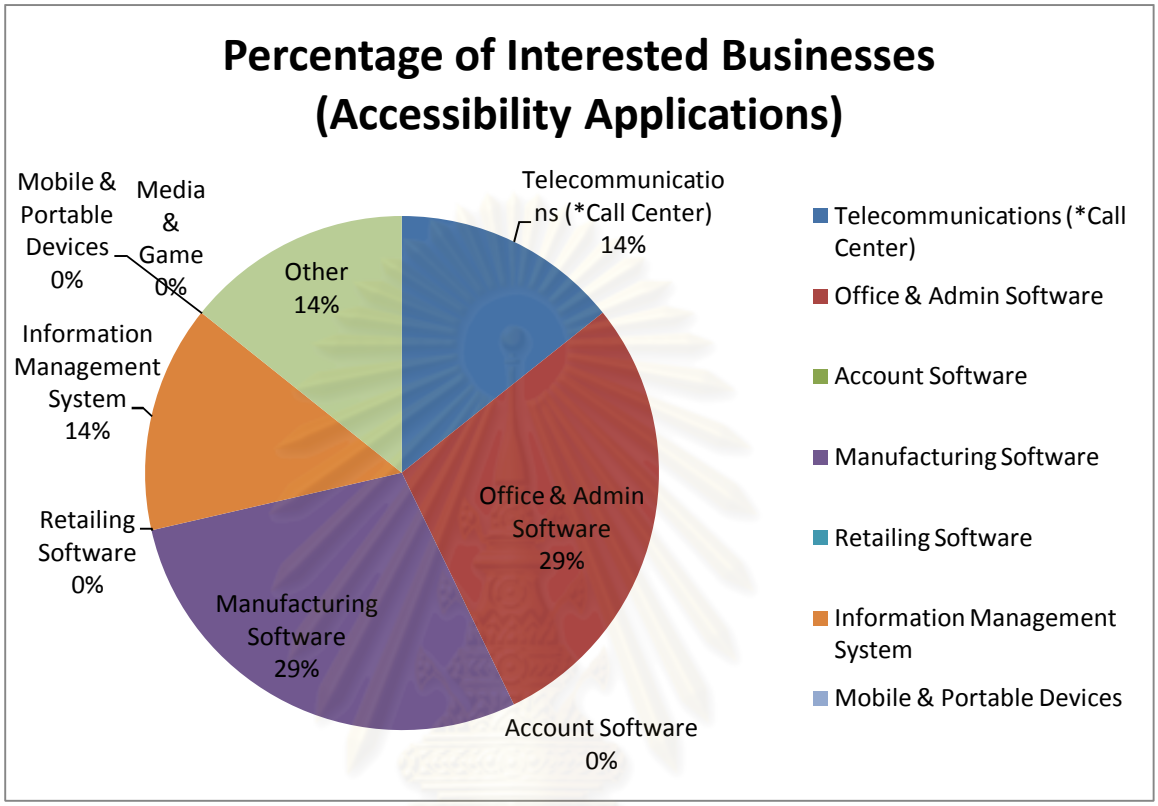


Figure 75 Percentage of Interested Businesses (Accessibility Applications)

Most accessibility applications are embedded applications that help support the core function and increase value. The details of the some suggested accessibility applications are as follows:

- Announcement by TTS synthesis voice – The announcement in organizations can be created by TTS system instead of using pre-recorded voice. It is suitable for dynamic information that changes frequently for example machine maintenance data in factories. This function is rather a support function but not the core function of manufacturing software.
- Create interaction channel with computer system by voice (TTS and ASR) – TTS and ASR system can be used to create interaction channel with systems such as administration and information system. Users can access to data more easily by

voice interface without the need of using keyboards. This can benefits mass users, which a lot of them do not know how to use computers or they are handicapped. Some examples of the plans are the administrative system for school that enables the parents to access to students' grades by using voice (ASR) and respond back with synthesis voice (TTS). Another example is the library searching system for blind people, which they can get access to the specific files they want by using voice.

- Enable communication by voice with webpage (TTS and ASR) – Voice can be used as an interaction medium with webpage that can create an innovative gimmick that adds value to the traditional webpage. Users can command webpage by voice interface without the need of using keyboards, and the system will respond to the user's command by synthesis voice.

6.1.2 Suggestions for Future Capabilities to Improve

TTS Capabilities

For accessibility applications, speech intelligibility is the most important aspect to improve since the information for accessibility applications is mostly specific information such as student's name, machine's series, book names, abbreviations, etc. NECTEC must increase the accuracy of the TTS system for the specific names and the abbreviations that sometimes are misread.

However, the naturalness of the TTS system matters less than with the call center application since most of the announcements or information read are short and concise. NECTEC's TTS system is considered acceptable to be commercialized at the present with some improvements on the specific names.

ASR Capabilities

For accessibility applications, the accuracy and the working environment are the most important concerns. The accuracy of NECTEC's ASR system must be improved since most of the input data is mostly specific information such as student's name, machine's series, book names, abbreviations, etc. For these specific names, the error

rate tends to be higher than general words. Therefore, in order to meet the required level of accuracy, the specific sets of words for different industries have to be created in the database. This might requires more effort from NECTEC in order to target the accessibility applications since it is applicable for wide ranging industries.

As for the working environment, the suitable working environment for accessibility applications is office to outdoor environment because some applications might need to be at outdoor environment such as schools or university.

The current domain of isolated words is sufficient for the accessibility application because mostly the inputs are isolated words.

Business Aspects

The main issue that is an obstacle for NECTEC to commercialize speech technologies with accessibility applications is that the applications are wide ranging and can be used with various industries. This makes it hard to target a specific industry, not like the call center application that targets the telecommunications sectors directly. Therefore, the customers are not likely to have much knowledge in this field since their core competencies are not related to speech. In order to commercialize speech technologies with accessibility applications, the increase in marketing and knowledge transfer are required. NECTEC could potentially suggest the applications for the customers since most customers might not be able to see the potential of speech technologies with their core business.

6.2 Second Future Target: Mobile Applications & E-book Reader

6.2.1 Details for Mobile Applications & E-book Reader

Mobile applications and e-book reader market are other interesting sectors for speech technologies. For mobile applications, the nature of the device is suitable for using speech technologies since it is related to speech. As for e-book reader, speech can be an interesting value-added function for the device. However, for both of the applications, due to the technological readiness of the speech technologies of NECTEC

that are not suitable with mobile platform at the present, the mobile applications and e-book reader can be an interesting future market.

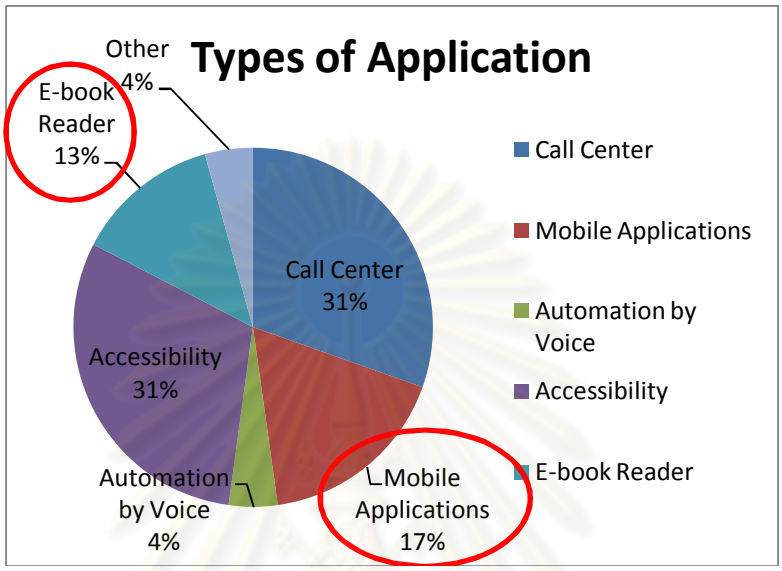


Figure 76 Types of Application

Speech technologies can help support the core function of mobile device and e-book reader to create gimmicks and add values. The details of the some suggested mobile applications and e-book reader applications are as follows:

- Voice command on mobile device (TTS and ASR) - Functions in mobile phones can be controlled by voice without the need of keypads such as voice dialing.
- E-book reader by voice (TTS) – As a value-added function to the traditional e-book reader, users can listen to the book instead of reading by themselves.

6.1.2 Suggestions for Future Capabilities to Improve

Technical Limitation

Currently, the speech technologies of NECTEC (both TTS and ASR) are not ready to be used in the mobile platform yet due to the required memory size. NECTEC's TTS technology will be able to support the mobile platform by approximately next year, which will be able to implement with mobile applications and e-book readers.

TTS Capabilities

Similarly to the call center applications, speech quality, naturalness, and speech intelligibility are all important aspects for mobile applications since the users will expect the level of quality to be comparable to the pre-recorded voice.

The speech quality is already in the sufficient level. However, for naturalness and speech intelligibility aspect, the improvement on the naturalness and accuracy should be increased since the users will expect high level of quality that is comparable with pre-recorded voice.

ASR Capabilities

Similarly to the call center applications, the accuracy and the working environment are the most important concerns. The accuracy of NECTEC's ASR system should be increased in order to match with the demand of the customers. As well as the working environment, it should be able to be used at any environment; therefore, the capability level should be increased from office environment to outdoor environment. However, for the current domain of isolated words is sufficient for the mobile applications because mostly the voice commands are isolated words.

Business Aspects

The mobile applications are interesting applications for speech technologies since the nature of the device is related to voice. Speech technologies for mobile devices will be increasing in the near future and the mobile market is an interesting market to enter. According to LG Mobile's 2011 Business Direction press conference, the mobile market worth in 2010 is 37,000 million THB (approximately US$ 1,150 million) with an expected growth rate in 2011. (LG Mobile's 2011 Business Direction, cited in www.pctodaythailand.com) Therefore, NECTEC should have the software for mobile platform ready in the near future.

For e-book reader device, speech technologies will be one value-added function that customers can look for. In the near future, the e-book reader will be increasing in numbers due to the lower price of the e-book reader. Therefore, speech technologies can be one competitive value-added feature.

6.2 Future Approximate Market Worth That NECTEC Could Potentially Gain (Mobile Application)

In the future, NECTEC could possibly gain revenue from accessibility application and mobile application. However, since accessibility application can be used in very wide-ranging businesses, it is not possible to estimate the market worth at the present. However, the mobile application market worth can be estimated as follows. Mobile application is another interesting prospective market, which NECTEC should target in the future when the technical capability is ready.

According to Opus Research, the market worth of global speech technology in mobile business was approximately US$ 125 million. (Opus Research, cited in Business 2.0 Magazine, 2007)

According to IDC research, the global market worth for mobile application in 2005 was estimated to be US$ 1.2 billion. (IDC, cited in Business Wire, 2006) According to Compass Intelligence, the global market worth for mobile application in 2007 was estimated to be US$ 3.8 billion. (Compass Intelligence, cited in TMCnet's article by Prabhala Ranga Sai, 2007) From the two year data, it can be estimated that the global market worth for mobile application in 2006 was approximately US$ 2.5 billion. This makes speech technology in the mobile business worth 5% out of the total mobile application market worth (US$ 125 million to US$ 2.5 billion). This can be summarized in the following figure:

Figure 77 Global Speech Technology in Mobile Application Market Worth

According to NECTEC's "Thailand ICT Market 2009 and Outlook 2010" report, the mobile application market worth in 2009 is estimated to be 3.069 billion THB. And with the estimation of the growth rate of 21.2%, NECTEC forecasted that the mobile application market worth in 2010 is 3.720 billion THB. (NECTEC, 2010) Using the growth rate of 21.2%, the mobile application market worth is estimated to be 4.51 billion THB in 2011 and 5.46 billion THB in 2012.

Assuming NECTEC will begin to seriously commercialize speech technology in 2011. Using to 5% figure from the estimation of the world's speech technology proportion in mobile business (see above) to forecast the market worth of speech technology in mobile business in Thailand in the year 2011, it could be forecasted that the market worth of speech technology in mobile business in 2011 is about 225 million THB. Assuming the market worth of the two technologies (TTS and ASR) are divided equally, 112.5 million THB are contributed to TTS and ASR technology each. Using the approximation of the market share of NECTEC (approximated by the experienced NECTEC's researchers), the share of NECTEC's TTS technology is approximately 30%, which is 33.75 million THB. And the share of NECTEC's ASR technology is approximately 20%, which is 22.5 million THB. **The approximated possible market worth**

that NECTEC could potentially gain in 2011 is about 56.25 million THB. This is summarized in the following figure:

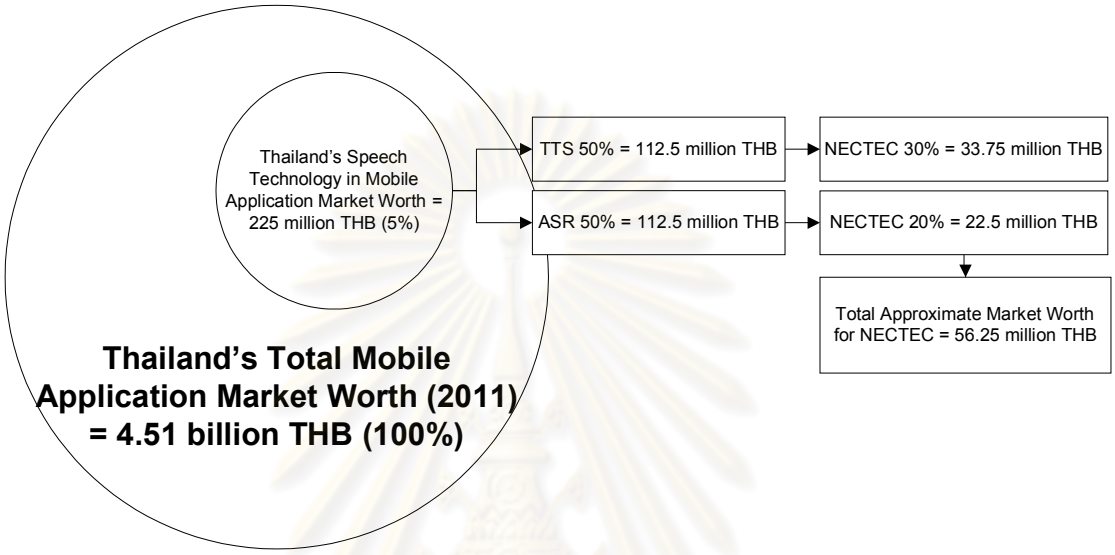


Figure 78 Thailand Speech Technology in Mobile Application Market Worth

However, if in the future NECTEC could improve the capabilities of the software and increase marketing and technology transfer, more share could potentially be gained. The scenario of NECTEC gaining 50% share in both TTS and ASR technology in year 2012 can be summarized as follows with the potential market worth of 136.5 million THB.

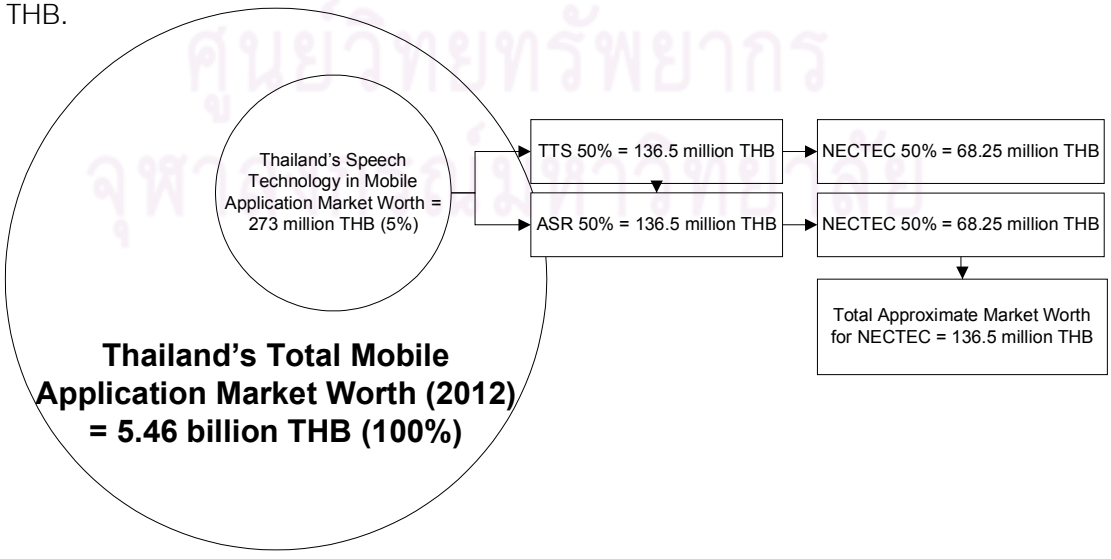


Figure 79 Thailand Speech Technology in Mobile Application Market Worth (Another Scenario)

Chapter VII

Conclusion and Suggestions

Speech technologies, in particular Text-to-Speech (TTS) and Automatic Speech Recognition (ASR), have been developed by NECTEC for a long time and will be ready to be commercialized in the near future. The Thai language speech technologies have made the progress in the research and development slowly comparing to other language such as English due to the complexity of its nature. Speech technologies are not commonly known among the businesses in Thailand, but the technologies have been successfully implemented in many other countries such as the United States, the United Kingdom, and Japan for a long time already. Currently the TTS technology software, VAJA 6.0, has been confirmed by NECTEC of the efficiency and effectiveness that it is ready to be commercialized. As for the ASR technology software, iSpeech, its readiness for commercialization is expected to be in the next year or two. For ASR technology that is planned to be commercialized, the level of its capability is limited to limit domain, which can input only isolated words.

The objective of the thesis is to suggest the applications of Thai speech technologies (TTS and ASR) that are viable in market aspect to be commercialized in Thailand's business environment. The thesis aims to identify the potential market of the applications of existing speech technologies in Thailand, and study the future potential market for further technology development.

Basically, the study was conducted based on two main areas: the study of the applications in other countries, and conducting the interviews with the potential customers of NECTEC. The potential customers of NECTEC are the software houses, which will be interviewed about their interest and plan to implement speech technologies with their software. The first interview was conducted with the software houses in various industries to identify the most probable group of customer. This interview suggested that the most prospective group of software is the call center software due high interest rate, the suitable nature of the call center with speech

technologies, the attractive growth rate of call center business in Thailand, the successful examples of international call center applications, the calls for value-added function for call centers, and the similar nature among the call center software developers which makes the targeting easy. The second interview was an in-depth interview conducted with the call center software developers to identify the plan of applications for speech technologies in call center. Furthermore, the interview with the customers' of the software houses, which are the end users of the call center solution, was conducted to confirm the demand of speech technologies for call center applications.

It can be summarized that the current most potential group of customer to implement speech technologies with is the call center software. The services that these developers plan to implement speech technologies with are the support to the traditional IVR system. General tasks can be taken over by speech technology system to reduce the need of human operators for simple and repetitive tasks. TTS system is suggested to be used as a replacement to the pre-recording voice that can provide repetitive, dynamic, and frequently changing information. ASR system is suggested to be used as an alternative to the traditional touchtone system (keypads). The plans to implement ASR system are to take input by voice such as menu bypassing, voice search, call routing, and input by voice. During the initial phrase of implementation, the choices between inputting by voice (ASR system) or by keypads (touchtone system) should be given.

According to the analysis, the future customer groups that NECTEC should focus are the accessibility applications and the mobile applications & e-book readers. Accessibility application has very wide ranging solutions and can be used in a wide variety of businesses, but the drawback is that it would be hard to focus on a particular group of customer. In the future with more technology and knowledge transfer, the interest rate could be increased by giving the potential customers more information. For mobile applications, the nature of the device is suitable for using speech technologies

since it is related to speech. As for e-book reader, speech can be an interesting value-added function for the device.

7.1 Study of International Applications

Currently, due to the increasingly higher computer performance and sophisticated basic speech technology, the speech applications are replacing the old touchtone system.

The capability of current state-of-the-art ASR system is able to understand complete continuous sentences and also extract information from them. As for TTS system, the technological advancement in the past few years made the naturalness of the speech increased significantly that it is very difficult to tell the difference between the speech synthesizer and a human. In the past few years the techniques of speech technologies have advanced significantly. The innovations in speech applications are constantly increasing, mainly in the telephone-based applications, and also other areas as well. The users accepted speech technologies as an accepted standard. Some successful applications of Nuance Communications have been elaborated in the previous section including:

- Accessibility
- Automation by Voice
- Automotive
- Call Center
- Mobile Applications
- Employee Productivity
- Transcription

All suggested areas are possible for Thailand except the transcription area, which the capability level of ASR cannot support the application yet.

7.2 Market Study of Current Potential Market

7.2.1 Business Environment of Thailand

The speech technologies are considered new technology for Thailand, and high investment cost must be covered for the research. In Thailand, only several sectors are involved in developing and commercializing Thai language speech technologies in basic technology form.

From SWOT Analysis and Porter's Five Forces Analysis, the basic speech technology industry is considered attractive, but the sectors that want to enter this industry must be large enough in size, power, and capital investment to initiate adequate research and development.

7.2.2 NECTEC's Current Capabilities

VAJA 6.0 is Thai language speech synthesis software developed by NECTEC. The technology enables the conversion of Thai text into synthetic speech that people can understand. The capabilities level of VAJA 6.0 can be summarized as follows:

Speech Quality	High quality
Naturalness	Similar to human's speech
Speech Intelligibility	Approximately 97% accuracy
Domain	Any word in Thai language including words that are not found in dictionary.
Other Capabilities	Input specific words such as name to authorize the pronunciation.

iSpeech software is speech recognition system that can recognize the Thai language speech. The system converts spoken words into parameters that it can understand. The iSpeech system can be categorized by the level of domain that it can recognize. There are three main types of iSpeech:

- iSpeech- W: Isolated word recognition (IWR)
- iSpeech- R: Continuous speech recognition (CSR) with limited grammar

- iSpeech- N: Continuous speech recognition (CSR) with unlimited grammar

Currently, only iSpeech- W is ready in terms of technical capabilities to be commercialized. iSpeech- W is the ASR system that can recognize isolated words. The capabilities of iSpeech- W can be summarized as follows:

Accuracy	More than 90% accuracy when used to recognize 100 words in an office environment.
Domain	Isolated words
Working Environment	Quiet room to office environment
Number of Words Able to be Recognized	Can input the unlimited number of words, but the accuracy is also decreased proportionally.

7.2.3 Market Segmentation and Targeting

From the market segmentation and analysis, the most prospective market segment for speech technologies in Thailand at present is the call center solution. It can be seen that the interview suggested the highest interest rate. Other than the demand from the call center business other reasons that support the targeting of call center business in the first stage are:

- The nature of the call center business that speech is suitable for speech technologies.
- The attractive growth rate of call center business in Thailand.
- Successful international call center applications commercialized.
- The calls for the value-added functions for call centers.
- The similarity in nature among the call center group that makes the targeting easy.

7.3 Targeting the Call Center Business

7.3.1 Overview of Call Center Business in Thailand

Nowadays, telephone is still the most heavily used communication channel due to the convenience and the reasonable cost. The call center business in Thailand can be segmented into three main service sectors: outsourced call center service, call center system and software service, and telemarketing service. In this study, the focus is mainly on the call center software service. The market worth and the growth of the call center business in Thailand suggested the opportunities for speech technologies to be implemented in Thailand, which from the previous section suggested that the market worth of the call center software business is approximately 595 million THB (2009). From the approximation, in 2011, NECTEC could potentially gain 27 million THB from this market. Also, the competition in the call center business tends to increase in the upcoming years, which calls for the call center software developers to increase their competitive advantages by speech technologies.

7.3.2 Market Study of Call Center Business

From the interview, almost all of the call center software developers have interest in speech technologies and have the plan to implement speech technologies with their current applications. The services that these developers plan to implement speech technologies are the support to the tradition IVR system. General tasks can be taken over by speech technology system and reduce the need of human operators for simple and repetitive tasks. TTS system is suggested to be used as a replacement to the pre-recording voice that can provide repetitive, dynamic, and frequently changing information.

ASR system is suggested to be used as an alternative to the traditional touchtone system (keypads). During the initial phrase of implementation, the choices between inputting by voice (ASR system) or by keypads (touchtone system) should be given since ASR system has rarely been implemented in Thailand. The plans to implement ASR system are to take input by voice such as menu bypassing, voice search, call routing, and input by voice.

The interview asked about the capabilities of TTS and ASR software for overall speech technologies in Thailand and for NECTEC's software. From the interview, NECTEC's TTS software (VAJA 6.0) has the potential to be commercialized at present. The focus might be put on improving the naturalness and speech intelligibility to ultimately increase the customer satisfaction. As for ASR, most customers still doubt about the accuracy issue of NECTEC's system (iSpeech). At present, NECTEC should improve the accuracy for the system with the current domain level of isolated words. Also, for the working environment that ASR system can be used with, NECTEC should improve from office level to outdoor level. Then if both of the issues suggested are improved, efforts can then be moved to improve the domain level to the sentence level in the future.

The interview on the end users could potentially suggest that the end users have interest on speech technologies. From all the study on the call center business, it can be suggested that the speech technologies have a promising future in the call center business, which is growing and still have rooms for a lot of improvements.

7.4 Market Study of Future Potential Market

The interesting future markets for speech technologies in Thailand are the accessibility applications and the mobile applications & e-book readers.

7.4.1 Accessibility Applications

An accessibility application has the highest interest rate along with the call center application. However, accessibility application has very wide ranging solutions and can be used in a wide variety of businesses, but the drawback is that it would be hard to focus on a particular group of customer. In the future with more technology and knowledge transfer, the interest rate could be increased by giving the potential customers more information. NECTEC could potentially suggest the applications for the customers since most customers might not be able to see the potential of speech technologies with their core business. Most accessibility applications are embedded applications that help support the core function and increase value by increasing

accessibility of information among people. Some suggested accessibility applications are as announcement by TTS synthesis voice, create interaction channel with computer system by voice, and enable communication by voice with webpage.

For accessibility applications, speech intelligibility is the most important aspect to improve since the information for accessibility applications is mostly specific information such as names, machine's series, book names, abbreviations, etc. NECTEC's TTS system is considered acceptable to be commercialized at the present with some improvements on the specific names. The accuracy of NECTEC's ASR system must be improved since most of the input data is mostly specific information. As for the working environment, the suitable working environment for accessibility applications is office to outdoor environment because some applications might need to be at outdoor environment.

7.4.2 Mobile Applications & E-Book Readers

For mobile applications, the nature of the device is suitable for using speech technologies since it is related to speech. As for e-book reader, speech can be an interesting value-added function for the device. However, for both of the applications, due to the technological readiness of the speech technologies of NECTEC that are not suitable with mobile platform at the present, the mobile applications and e-book reader can be an interesting future market. Some suggested mobile applications and e-book reader applications are voice command on mobile and e-book reader by voice.

Currently, the speech quality is already in the sufficient level. However, for naturalness and speech intelligibility aspect, the improvement on the naturalness and accuracy should be increased since the users will expect high level of quality that is comparable with pre-recorded voice. For ASR, the accuracy and the working environment are the most important concerns. The accuracy of NECTEC's ASR system should be increased in order to match with the demand of the customers. As well as the working environment, the capability level should be increased from office environment to

outdoor environment. However, for the current domain of isolated words is sufficient for the mobile applications because mostly the voice commands are isolated words.

According to NECTEC, the market worth of the mobile application business is approximately 3.069 billion THB (2009). From the approximation, in 2011, NECTEC could potentially gain 56.25 million THB from this market.

7.5 Suggestions for NECTEC

At present, NECTEC should focus on targeting the call center software developers with improvements on the suggested areas elaborated above. NECTEC should place importance on customer relationship management with the interested call center companies that already have plans in their minds and provide them with both the technical and business supports on implementations. As for the companies that are interested but still lack of knowledge and plan in speech technologies, supports should be provided to them in order to help them with the implementations. This can gain NECTEC a strong customer base. The suitable licensing plan should be created and adjusted according to the needs of the customers. Also, clear understanding on the expectations of the technologies should be created between NECTEC and the customers.

Along with targeting the call center business, NECTEC should increase the market awareness in speech technologies to businesses in Thailand in order to increase the interest rate. More marketing effort and technology transfer can increase the awareness of this emerging technology among the businesses in Thailand. In this way, NECTEC can potentially gain in terms of both increasing the customer base and providing support to the industry (one objective of NECTEC to help support and increase knowledge for industries in Thailand). Trustworthy researches that support the investment of speech technologies can be provided to help the customers make decisions. For the future, the possible future market can be targeted at the accessibility applications and the mobile applications & e-book readers. NECTEC could potentially fix the current technical limitations suggested above in order to target those groups.

In summary, this study provides a guideline for NECTEC and other interesting parties for the potential business opportunities of speech technologies (TTS and ASR) in Thailand's business environment. For the actual implementation, further in-depth business feasibility study can be carried out in order to support the implementations.

References

Artelt, D. **Voice Compass: International 2008/2009: Speech Goes Mainstream**. aixvox GmbH, 2008

Bergevin, R., Kinder, A., Siegel, W., and Simpson, B. **Call Centers for Dummies: 2nd Edition**. John Wiley & Sons Canada, Ltd., 2010

Borzo, J. Now You're Talking. **Business 2.0 Magazine** [Online]. 2007. Available from : money.cnn.com [2007, March]

Business Wire. **Worldwide Mobile Enterprise Application Market Will Experience Double-Digit Growth and Reach $3.5 Billion by 2010, IDC Reveals** [Online]. 2006. Available from : www.businesswire.com [2006, December]

Chai Wutiwiwatchai, and Furui, S. **Thai Speech Processing Technology: A Review, published in Speech Communication. Elsevier B.V.**, 2006

Clifton, D. S. Jr., and Fyffe, D. E. **Project Feasibility Study Analysis: A Guide to Profitable New Ventures**. John Wiley & Sons, Inc., 1977

Compass Intelligence [Online]. Available from : www.compassintelligence.com

Datamonitor [Online]. Available from : www.datamonitor.com

Dutoit, T. An **Introduction to Text-To-Speech Synthesis**. Kluwer Academic Publishers, 1997.

Eastwood, G. **The Voice Business Market Outlook: Key Technologies and Emerging Market Opportunities**. Business Insights, Ltd., 2005

Frost & Sullivan. **Thailand Contact Center Market: Market Insight**, 2008

Ferrell, O. C., and Hartline, M. D. **Marketing Strategy**. Cengage Learning, 2008

Furness, V. **Call Center Growth Strategies: Vendor opportunities in North America and Europe**, Business Insights

Global Industry Analyst Inc. **Speech Technology: A Global Strategic Business Report**, 2011

Israel, G. D., **Determining Sample Size**, University of Florida IFAS Extension

International Data Corporation (IDC) [Online]. Available from : www.idc.com

Johnston, D., Whittaker S. J., and Attwater, D. J. **An Overview of Speech Technology for Telecom Services in the United Kingdom**. BT Laboratories, 1996

Kasikorn Research Center [Online]. Available from : www.kasikornresearch.com

LG. **LG Mobile's 2011 Business Direction**, cited in www.pctodaythailand.com , 2011

Meredith, J. R., and Mantel S. J. Jr. Project **Management: A Managerial Approach 6th Edition**, 2006

National Electronics and Computer Technology Center (NECTEC), 2010

National Electronics and Computer Technology Center (NECTEC). **Thailand ICT Market 2009 and Outlook 2010**, 2009

Nuance Communications [Online]. Available from : www.nuance.com

Office of Small and Medium Enterprises Promotion [Online]. Available from : www.sme.go.th

Opus Research [Online]. Available from : www.opusresearch.net

Porter, M. E. **On Competition**. Harvard Business School Publishing Corporation, 2008

Sai, P. R. **US Mobile Applications Market Expected to Reach $9 Billion By 2011** [Online]. 2007. Available from : www.tmcnet.com [2007, March]

Sorawit Buasri. **Thailand Contact Center 2009**. Contacts Mind Co.,Ltd, 2009

Thailand's Department of Business Development [Online]. Available from : www.dbd.go.th

The Federal of Thai Industries [Online]. Available from : www.fti.or.th

The Nation Newspaper [Online]. Available from : www.nationmultimedia.com

Warwick Manufacturing Group (WMG). **Technology Management Module Note**, 2009

Young, S. **The HTK Book (for HTK version 3.4)**. Cambridge University Engineering Department, 2006

Appendices

Appendix A

Questionnaire on "Market Study of Thai Speech Technologies in Thailand"

The objective of this questionnaire is to study the business opportunities of implementing speech technologies with businesses in Thailand. This questionnaire aims to interview the software developers that have potential to license speech technologies and develop into applications. This questionnaire focuses on the Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) technology.

PART A : Basic Information of the Company

A1 : Name / Address / Contact Person

Please provide the following information on your company

Company's Name (Thai) .....

Company's Name (English) .....

Contact Person Name : ..... Last Name : .....

Position : ..... E – mail : .....

Name of Managing Director : ..... Last Name: .....

Headquarter Address NO : ..... Building : ..... Floor : .....

Soi : ..... Road : .....

Sub-District : ..... District : .....

Province : ..... Postal code : .....

Telephone No. : ..... Fax No. : .....

E – mail : ..... Website : .....

Year of establishment : .....

A2: Product / Services of the Company

1.1 What type of software business are you operating in? (More than 1 answer is applicable)

❑ 1. Developer
❑ 2. Importer/Sale Representative
❑ 3. Distributor
❑ 4. Retailer
❑ 5. Computer Services
❑ 6. Other (Please specify.....)

1.2 Which type of software do you develop or provide services in?

❑ 1. Enterprise Software
❑ 1.1 Package Software ❑ 1.2 Outsource Software Development ❑ 1.3 Other (Please specify.....)

❑ 2. Mobile Applications

❑ 2.1 Business Application — O 2.1.1 Package Software — O 2.1.2 Outsource Software Development

❑ 2.2 Entertainment Application — O 2.2.1 Package Software — O 2.2.2 Outsource Software Development

❑ 2.3 Other (Please specify .....)

❑ 3. Embedded Software

❑ 3.1 Package Software — ❑ 3.2 Outsource Software Development — ❑ 3.3 Other (Please specify.....)

❑ 4. Other (Please specify .....)

2. Which sectors are your customers in? Please specify the proportion. (More than 1 answer is applicable

Customer Sectors	Year 2010 (%)
1. Public Enterprise	
2. Private Enterprise (Continue with 2.1)	
3. Consumer & Small Enterprise	
Total	**100**

2.1 What are the sizes of the private enterprises? Please specify the proportion.

Only for Private Enterprise	Year 2010 (%)
Medium Size Private Enterprise*	
Large Size Private Enterprise*	
Total	**100**

NOTE: *Medium size enterprise has 20-100 computers.
**Large size enterprise has more than 100 computers.

3. Which industries are your customers operating in? Please specify the proportion. (More than 1 answer is applicable)

Customer Industries	Year 2010 (%)
1. Financial & Banking (Banks, Insurance, Financial Institutes)	
2. Travel & Leisure (Hotels, Restaurants, Tour Agencies)	
3. Logistics	
4. Medical & Health Care	
5. Educations (Schools, Universities)	
6. Agricultural	
7. Automotive Manufacturing	
8. Manufacturing (Other)	
9. Telecommunications	
10. Retailing	
11. Other	
Total	**100**

4. Which platform does your company have competencies in? (More than 1 answer is applicable)

❑ 1. Unix — ❑ 2. Mobile — ❑ 3. Open source
❑ 4. RFID, Embedded Software — ❑ 5. Windows — ❑ 6. Other (Please

Specify.....)

A3 : Human Resource of the Company

5. Please specify the number of employees in your company.

Employees	Year 2010
	Number
1. Technical Employees	
2. Management and Non-technical Employees	
Total	

PART B : Interest & Awareness in Speech Technologies

6. Do you have interest to implement speech technologies with your current business?

- ❑ Never heard of speech technologies — Briefly explain the technologies + Skip to PART E
- ❑ Aware of speech technologies but do not have interest to use with the business — Skip to PART D
- ❑ Aware of speech technologies and is moderately interested (Interested to use the technologies but might not have specific plan yet) — Continue with PART C
- ❑ Aware of speech technologies and is very interested (Have a plan to implement speech technologies in their mind) — Continue with PART C

7. What do you think is the future direction of speech technologies in Thailand?

- ❑ Decreasing _____
- ❑ Constant _____
- ❑ Increasing Slowly _____
- ❑ Increasing Sharply _____

PART C : Plan to Implement Speech Technologies

(Only for Companies that are Moderately and Very Interested)

7. Which type of application(s) do you plan to use speech technologies with? (More than 1 answer is applicable)

- ❑ Application on PC
- ❑ Mobile Application
- ❑ Solution

❑ Embedded Application

Please briefly explain about the application. _____
_____
_____

9. Which technology do you plan to implement? (More than 1 answer is applicable)

❑ Text-to-Speech (TTS)
❑ Automatic Speech Recognition (ASR)
❑ Automatic Speaker Verification (ASV)

10. The customers of the application will be from new market or from existing market?

❑ New Market Go to 12.1
❑ Existing Market Go to 12.2

10.1 For the new market, will there be any barrier to entry?

❑ Yes, please specify _____
❑ No

10.2 For the existing market, will there be any change in marketing strategy?

❑ Yes, please specify _____
❑ No

11. What are the main objectives of implementing speech technologies with your business?

❑ Add value to the existing products / services
❑ Increase efficiency of processes
❑ Increase customer satisfaction
❑ Create innovation
❑ Use as a core application
❑ Other, please specify _____

12. What are the benefits from this application?
_____
_____
_____
_____
_____

PART D : Problem in Implementing Speech Technologies

13. Do you have any experience on implementing speech technologies?

❑ Yes Go to 15.1
❑ No Go to 15.2

13.1 What are some problems you have already encountered?

Technological Aspect

_____
_____
_____
_____
_____

Business Aspect

_____
_____
_____
_____
_____

13.2 What are some problems you expect could happen?

Technological Aspect

_____
_____
_____
_____
_____

Business Aspect

_____
_____
_____
_____
_____

PART E : Expectations from Implementing Speech Technologies

**Only for the ones who never heard of speech technologies will be briefly explained about the speech technologies and continue with number 14.*

14. After you have been briefly explained about speech technologies, do you think you have interest in implementing speech technologies with your business?

❑ No interest
❑ Interested and need further information to make further decisions

15. What is your expectation of speech technologies? (More than 1 answer is applicable)

- ❑ Speech technologies will be widely used in Thailand
- ❑ More clarity on how speech technologies can be applied to business
- ❑ Increase in the technologies' efficiency and effectiveness
- ❑ Other _____

PART F : Explanations

Speech Technologies	Explanations
Text-to-Speech (TTS)	The conversion of text into voice speech by separating, compiling, and generating the voice. This is different to the pre-recording speech that the dynamic information that changes frequently can be read such as names. The advantages of using this technology are convenient and eyes-free.
Automatic Speech Recognition (ASR)	The conversion of voice speech into text by receiving voice input and compiling by probability models. Some examples are the applications that require repetitive input such as credit number. The advantages of using this technology are convenient, eyes-free, hands-free, which require no keyboards or keypads.

Questionnaire on "Market Study of Thai Speech Technologies in Thailand"

The objective of this questionnaire is to study the business opportunities of implementing speech technologies with businesses in Thailand. This questionnaire aims to interview the software developers that have potential to license speech technologies and develop into applications. This questionnaire focuses on the Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) technology.

Company's Name (Thai) .....

Company's Name (English) .....

PART A : Applications of Speech Technologies

1. Which group of customer do you think will be suitable to implement speech technologies' applications with? (More than 1 answer is applicable)

Customer Industries	Tick if Suitable
1. Financial & Banking (Banks, Insurance, Financial Institutes)	
2. Travel & Leisure (Hotels, Restaurants, Tour Agencies)	
3. Logistics	
4. Medical & Health Care	
5. Educations (Schools, Universities)	
6. Agricultural	
7. Automotive Manufacturing	
8. Manufacturing (Other)	
9. Telecommunications	
10. Retailing	
11. Other	
12. Any group	

2. Why do you think those groups of customer are suitable with speech technologies?

_____
_____
_____
_____

3. Please explain briefly about your plan of speech technologies' application.

_____
_____
_____
_____

PART B : Revenue / Licensing / Marketing Strategy

4. Revenue & Market Share

4.1 Company's Revenue Per Year =
_____

4.2 Company's Market Share =
_____

5. Which type of licensing agreement with NECTEC do you think is the most appropriate?
   (For example pay one large upfront fee or pay smaller revenue shares)

_____
_____
_____
_____

6. What marketing strategy do you plan to use with your customers?
   ❑ Bundle with existing packages and increase the selling price
   ❑ Provide as an option to the existing packages and the customers pay for the extra fees
   ❑ Provide as a value-added function with no extra changes in order to increase customer satisfaction
   ❑ Other, please specify

_____

PART C : Technical Capabilities of Speech Technologies

7. Do you think the current level of speech technologies is sufficient?

Text-to-Speech (VAJA 6.0)	❑ Have experience with NECTEC's VAJA 6.0 ❑ Never have experience with NECTEC's VAJA 6.0	
Capabilities	**Level of Capability**	**If not sufficient, which aspect can be improved?**
Speech Quality	❑ Sufficient ❑ Insufficient	
Naturalness	❑ Sufficient ❑ Insufficient	
Speech Intelligibility	❑ Sufficient ❑ Insufficient	

Automatic Speech Recognition (iSpeech)	❑ Have experience with NECTEC's iSpeech ❑ Never have experience with NECTEC's iSpeech	
Capabilities	**Level of Capability**	**If not sufficient, which aspect can be improved?**
Accuracy	❑ Sufficient ❑ Insufficient	
Domain Level	Expected domain level ❑ Isolated words ❑ Phrases/Sentences ❑ Continuous Conversation	
Robustness	Expected level of environment to be used with ❑ Quiet Room ❑ Office ❑ Outdoor	

PART D : Explanations

Speech Technologies	Explanations
Text-to-Speech (TTS)	The conversion of text into voice speech by separating, compiling, and generating the voice. This is different to the pre-recording speech that the dynamic information that changes frequently can be read such as names. The advantages of using this technology are convenient and eyes-free.
Automatic Speech Recognition (ASR)	The conversion of voice speech into text by receiving voice input and compiling by probability models. Some examples are the applications that require repetitive input such as credit number. The advantages of using this technology are convenient, eyes-free, hands-free, which require no keyboards or keypads.

Questionnaire on "Market Study of Thai Speech Technologies in Thailand"

The objective of this questionnaire is to study the business opportunities of implementing speech technologies with businesses in Thailand. This questionnaire aims to interview the end users of applications that have the potential to implement speech technologies. This questionnaire focuses on the Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) technology.

PART A : Basic Information of the Company

A1 : Name / Address / Contact Person

Please provide the following information on your company

Company's Name (Thai) .....

Company's Name (English) .....

Contact Person Name : ..... Last Name : .....

Position : ..... E – mail : .....

Name of Managing Director : ..... Last Name: .....

Headquarter Address NO : ..... Building : ..... Floor : .....

ที่อยู่ของสำนักงานใหญ่ Soi : ..... Road : .....

Sub-District : ..... District : .....

Province : ..... Postal code : .....

Telephone No. : ..... Fax No. : .....

E – mail : ..... Website : .....

Year of establishment : .....

A2 : Industry that the Company Operates in

1. Which sector does your company operates in?

❑ Public Enterprise

❑ Private Enterprise Size ❑ Small ❑ Medium ❑ Large

NOTE: According to Office of Small and Medium Enterprises Promotion, www.sme.go.th
Small size enterprise has less than 50 employees.
Medium size enterprise has 50-199 employees.
Large size enterprise has 200 or more employees.

2. Which industries do you operating in? Please specify the proportion. (More than 1 answer is applicable)

Customer Industries	Year 2010 (%)
1. Financial & Banking (Banks, Insurance, Financial Institutes)	
2. Travel & Leisure (Hotels, Restaurants, Tour Agencies)	
3. Logistics	
4. Medical & Health Care	
5. Educations (Schools, Universities)	
6. Agricultural	
7. Automotive Manufacturing	
8. Manufacturing (Other)	
9. Telecommunications	
10. Retailing	
11. Other	
Total	**100**

3. Please explain briefly about your business.

_____
_____
_____
_____
_____

PART B : Implementation of Speech Technologies with Call Center Function

4. Have you ever heard of speech technologies?

❑ Never heard of speech technologies (Briefly explain the technologies)
❑ Aware of speech technologies

Speech Technologies	Explanations
Text-to-Speech (TTS)	The conversion of text into voice speech by separating, compiling, and generating the voice. This is different to the pre-recording speech that the dynamic information that changes frequently can be read such as names. The advantages of using this technology are convenient and eyes-free.

Automatic Speech Recognition (ASR)	The conversion of voice speech into text by receiving voice input and compiling by probability models. Some examples are the applications that require repetitive input such as credit number. The advantages of using this technology are convenient, eyes-free, hands-free, which require no keyboards or keypads.

5. Do you have interest in implementing speech technologies with your call center function (Assuming no extra costs taken into account for now)

❑ No interest — Go to 5.1
❑ Interested — Go to 5.2 and 5.3

5.1 What are the reasons you do not want to implement speech technologies with your call center? (More than 1 answer is applicable)

❑ Customers may not be familiar with the new technology and turn away
❑ Do not have enough confidence in the technical reliability of speech technologies
❑ Concern that there might be extra indirect costs added
❑ Other, please specify _____

5.2 What do you think you will gain from implementing speech technologies with your call center? (More than 1 answer is applicable)

❑ Add value to the existing products / services
❑ Increase efficiency of processes
❑ Increase customer satisfaction
❑ Create innovative image
❑ Use as a core application
❑ Other, please specify _____

5.3 What is the acceptable range of extra costs for speech technologies comparing to the current costs you are paying?

_____
_____
_____
_____

Appendix B

	Company Name	Authorized Share Capital	Total Revenue (2009)	Call Center Business to Total Business (%)	Business Size	Total Revenue for Call Center (2009)	Outsourced?	Outsource to Total Business (%)
1		20,000,000		100	L			
2		3,000,000		100	M			
3		1,000,000		100	S			
4		50,000,000		100	L		Yes	50
5		2,000,000		100	S			
6		5,000,000		10	M			
7		5,000,000		100	M			
8		193,000,000		100	L		Yes	80
9		1,000,000		100	S			
10		1,000,000		100	S			
11		475,200,000		30	L			
12		1,000,000		10	S			
13		1,000,000		50	S			
14		1,000,000		10	S			
15		12,750,000		20	L			
16		1,000,000		8	S			
17		1,000,000		100	S			
18		3,000,000		100	M			
19		1,000,000		40	S			
20		1,000,000		100	S			
21		100,000,000		100	L		Yes	
22		6,000,000,000			L		Yes	
23		2,000,000			S			
24		1,000,000			S			
25		664,410,000		100	L		Yes	
26		10,000,000			M			
27		15,000,000			L			
28		2,000,000			S			

S : <3M				
M : >=3M to <=10 M				
L : > 10 M				
Average Revenue (S)	6,823,853		Market Value =	595,854,766
Average Revenue (M)	13,415,841			
Average Revenue (L)	48,137,957			
Frequency (S)	14			
Frequency (M)	5			
Frequency (L)	9			
Total Frequency	28			

Biography

Miss Puthita Wacharasin was born on June 9, 1987. In 2005, she finished her high school from Kasetsart University Laboratory School (International Program). In 2009, she graduated with a Bachelor of Engineering in Industrial Engineering, Chulalongkorn University (Thailand). In May of 2009, she enrolled as a part-time student of the Regional Centre of Manufacturing Systems Engineering for the Master of Engineering in Engineering Management from Chulalongkorn University (Thailand) and Master of Science in Engineering Business Management from the University of Warwick (United Kingdom).