# Language Science and Language Technology in Africa:

## A Festschrift for Justus C. Roux

Edited by

Hugues Steve NDINGA-KOUMBA-BINZA

&

Sonja E. BOSCH

SUN PRESS

CHAPTER 17

# VOICE USER INTERFACE DESIGN FOR EMERGING MULTILINGUAL MARKETS

Gerhard van Huyssteen[1], Aditi Sharma Grover[2], Karen Calteaux[2]

[1]*Centre for Text Technology, North-West University, Potchefstroom, South Africa*
gerhard.vanhuyssteen@nwu.ac.za
[2]*Meraka HLT Group, Council for Scientific and Industrial Research (CSIR)*
asharma1@csir.co.za, Kcalteaux@csir.co.za

## 1. INTRODUCTION

Socio-economic development in South Africa poses a variety of challenges, such as bridging communication gaps, providing access to information, dealing with pockets of technological and functional illiteracy, and addressing widespread poverty. South Africa is thus a prototypical example of a developing world nation – it remains marginalised in the world economy, and is considered a medium development country in terms of the United Nations Development Programme's (UNDP's) Human Development Index (International Marketing Council of South Africa, 2010). As a result, it offers many opportunities typical of an emerging region, including support for regional infrastructure projects and transferring knowledge and technology, aspects which are noted in the Economic Development in Africa Report 2010 as important avenues for promoting regional development (UNCTD 2010).

The country boasts a progressive Constitution that protects linguistic human rights and provides for the promotion and equitable use of all eleven of its official languages. A National Language Policy Framework adopted by the South African Government in 2003 mandates all government structures, and encourages private enterprises, to provide for the language needs of end-users. A right to access services in a language of choice is explicitly stated in the Framework (DAC 2003). The need for this is underscored by research indicating that there is not a single official language which can be used as a lingua franca across all linguistic communities in the country (PanSALB 2000).

Within this highly multilingual environment, telephones, and cell phones in particular, are a primary means of communication. In 2007 it was estimated that 87.6% of the South African population had access to a cell phone (Gapminder 2007/2008). By 2008 it was estimated that cell phone penetration in South Africa had increased to 92% (Gapminder 2007/2008) and by mid-2009 to 98% (IT News

Africa 2009). In stark contrast, it is estimated that only 9 to 11% of the South African population had access to the Internet in 2009[1]. One could therefore conclude that telephone-based services have the potential to bridge information and communication gaps prevalent in the South African society (Barnard *et al.* 2010) and other similar developing-world contexts (Tucker *et al.* 2004).

Based on the above, our point of departure for this research is that South Africa holds many market opportunities for multilingual (or multilingualised or multi-language) interactive voice response (IVR) systems, with which end-users could get access to information, complete transactions, activate services, make reservations, etc. This has also been the assumption of the African Speech Technology (AST) project (www.ast.sun.ac.za), which was at the time a ground-breaking human language technology (HLT) project in South Africa, led by Prof Justus Roux (Roux *et al.*, 2000). In subsequent years, various other research and development projects in South Africa were based on the very same assumption, including the Open Phone project (www.meraka.org.za/hlt_projects_ophone.htm) and the Lwazi project (www.meraka.org.za/lwazi), both led by the Meraka Institute of the Council for Scientific and Industrial Research (CSIR). All of these projects focussed by and large on the development of resources and technologies for automatic speech recognition (ASR) and text-to-speech (TTS) systems, as well as the development of prototypes of telephone-based services for all eleven official languages of South Africa.

As was pointed out in numerous publications from these and other projects, various interesting challenges arise when developing information and communication technologies (ICTs) for multilingual, emerging markets (e.g. Tucker *et al.* 2004; Barnard *et al.* 2008; Nasfors 2007). Designing effective voice user interfaces (VUIs) is even more challenging when very little is known about users' needs and/or preferences when accessing information via technologies such as IVRs (Barnard *et al.* 2008; Sharma Grover & Barnard 2011a). Previous research has indicated the potential of telephone-based services for stimulating socio-economic growth through access to information (Sharma Grover & Barnard 2011b), as well as some of the challenges associated with designing VUIs (Sharma Grover *et al.* 2009; Sharma Grover & Barnard 2011a; Plauché & Nallasamy 2007) for low-literate users, and taking into account dialectal differences and cost, respectively. Yet very few companies in South Africa specialise in multilingual IVR development; the commercial demand for such services seems to be limited; and very little research data is available in this domain. The VUI designer in commercial markets is therefore left in the dark – often having to base design decisions on own intuition.

The aim of our current research is to get a better understanding of business and design issues related to IVRs in a multilingual, emerging market such as South

Africa, in order to shed light on the challenges relating to VUI design for such markets. We were specifically interested in answering the following questions:

1.  What are the most important business drivers for implementing multilingual IVRs?
2.  If a multilingual IVR is to be implemented, how many languages should be available, and where/how should the language choice be offered?
3.  With regard to choice of voice (of the persona), what are current trends in South Africa?
4.  With regard to input modality (touch-tone or speech), why is speech not used more often?

Answers to these questions will provide a glimpse into the challenges facing VUI designers in the country, especially in terms of requirements analysis, a design framework and governance. In the next section, we explicate the methodology we followed to answer these questions. In section 3 we discuss the most important business drivers (and hurdles) related to the development of multilingual IVRs in South Africa. Section 4 focuses on some design issues, specifically related to language choice, options for choice of the voice, and input modality. The chapter concludes with a view on other issues and questions to be investigated in future research.

## 2.    METHODOLOGY

Our investigation into the multilingual IVR landscape of South Africa took a triangulated approach that included: (a) basic data collection on numerous South African telephone-based services; (b) interviews with IVR developers; and (c) interviews with companies that typically use IVR-based solutions for customer interaction and services.

We first conducted an impressionistic survey on a number of IVRs in South Africa by calling into these services to gather information relevant to our research questions. We were specifically interested in –

*   whether the service has a multilingual offering, and if so, which languages are offered, and where/how in the IVR the language choice is presented;
*   the characteristics (male/female; accent) of the voice artist used for the persona; and
*   whether the service uses speech (i.e. ASR) or touch-tone (i.e. dual-tone multi-frequency (DTMF) signalling) as input modality.

Since it is not always clearly discernable from a caller's perspective whether IVR or automatic call distribution (ACD) technology is being used, for our purposes we use the term "IVR" as an all-encompassing term to refer to technology that allows

humans to access computer systems (typically for self-service purposes, or to get access to automated information), or call/contact centre agents (or operators, consultants, representatives, etc.) through a telephone by using either speech or keypad input. We limited our survey to IVRs that required some form of input from the caller (e.g. to make a language choice, or to choose from a menu of options the kind of information or service s/he wants), and explicitly excluded services that merely route callers automatically to a contact centre or reception (i.e. "pure" ACDs). For our test battery, we did not make a distinction between call reasons – all systems were treated equally irrespective of whether the caller wants full self-service (e.g. telephone banking, or getting up-to-date flight information), information (e.g. how to apply for a social grant), or wants to speak to an agent (e.g. to change personal details).

We identified six domains for investigating IVRs: transport (airlines, airports, trains, etc.), banking, medical aid funds, telecommunications, government services, and entertainment (e.g. cinemas, satellite television, ticket services, etc.). Our choice of IVRs in each of these domains was meant to be exemplary rather than exhaustive; note that our choice was skewed by the fact that we tried to include at least all speech-based IVRs that we were aware of, and to also include at least a few services that we knew had a multilingual offering. One should also note that it is often quite challenging to gather accurate information for a survey like this, especially in commercial environments. For example, to assess an IVR offering telephone banking, the surveyor has to have an account with (or at least have access to dummy information for) that specific financial service provider before s/he can log on to that service to do an assessment. Since the aim of this survey was to get an impressionistic view of trends in designing IVRs in South Africa, we are not making any statistical claims regarding the IVR landscape in South Africa as a whole – such claims would be better left to exhaustive surveys. Table 1 presents an overview of the six domains and 34 IVRs we investigated.

**Table 1:** *Overview of domains*

| Domain | Examples | # of IVRs |
|---|---|---|
| Transport | Flight reservations, flight information, general enquiries | 7 |
| Banking | Telephone banking, general enquiries | 4 |
| Medical aid funds | Customer care, trauma help | 5 |
| Telecommunications | Pre-paid services, balance enquiries, general enquiries | 6 |
| Entertainment | Ticket reservation, service activation, general enquiries | 4 |
| Government services | Information access, reporting of crime or bad service | 8 |
| | **Total** | **34** |

The second part of our approach centred on gathering qualitative information through interviews with various role-players in the South African IVR and contact centre industries. VUI design services in South Africa are primarily provided by

companies operating in the contact centre, business process outsourcing (BPO) or ICT infrastructure industries, or companies that offer bespoke software solutions. These companies typically tend to provide clients with solutions that encompass services such as contact centre integration and automation, provision of hosting technology and infrastructure, contact centre operations, and automation of customer interactions with self-service applications. Thus, VUI design tends to become a smaller, ad hoc, value-added, service required to fulfil the larger solutions provided by these companies, as opposed to a specialised VUI design service often found elsewhere in the world.

We conducted detailed interviews with senior developers or managers at four such companies; of these one is a primarily state-owned entity that provides contact centre and BPO services to various government departments and related subsidiaries, another two are South African divisions of multi-national BPO companies providing comprehensive large-scale contact centre and IVR services (amongst other solutions), and the last one is a local South African small to medium sized enterprise (SME) that develops contact centre and IVR solutions.

Thirdly, we also conducted interviews with representatives from two companies in the banking and government services verticals; these companies are typical clients of the above-mentioned contact centre/IVR solutions provider companies. Both these representatives are responsible for the voice channel in their respective companies.

Through these interviews we sought to ascertain various issues and trends pertaining to multilingual IVR design and usage in South Africa. In addition to getting impressions relating to the questions above, we also sought information about the business drivers for multilingual IVR implementation.

For purposes of objectivity, we decided to treat all interviewees as anonymous; in the remainder of the study we will only refer to Respondent 1, 2, etc. We also keep the names of the IVRs called anonymous, since this chapter is not meant to promote or criticise any of the companies investigated.


## 3.    BUSINESS DRIVERS FOR MULTILINGUAL IVRs

Our survey of IVRs brought to light that multilingual IVRs are not as common as one would expect in a country where eleven official languages are constitutionally recognised; for example, in our survey of eight government services IVRs, only four offer service or information in all eleven languages. If one looks at commercial IVRs, it is clear that very few support more than English. In summary, of the 34 IVRs investigated, only nine have a multilingual offering. Given this view of the multilingual IVR landscape in South Africa, one could ask why multilingual IVRs are not more widespread, especially given the high premium placed on

295

multilingualism in the South African Constitution. What are the most important business drivers for (not) implementing multilingual IVRs? Elsewhere (Calteaux, Sharma Grover & Van Huyssteen 2012), we discuss these business drivers in more detail; suffice it to only mention some of these drivers here, including:

- Improved branding
- Cost saving (e.g. through increased automation)
- Revenue generation (by offering new or value-added services)
- Increased customer satisfaction
- Customer retention (i.e. preservation of existing revenue base)
- Customer delight
- Improved access to information or services
- Increased call centre agent morale (thereby reducing agent attrition)
- Increased agent utilisation
- Improved productivity
- Access to business intelligence for strategic advantage (e.g. call reasons, customer profiling, etc.)
- Opportunities for upselling of products
- Compliance with laws and regulations
- Political motivation (i.e. to have a multilingual offering could also serve a political agenda by (implicitly) subscribing to values of inclusiveness, nation-building, etc.)
- Competitive advantage (i.e. "If my competitor has a multilingual IVR, I should probably have one too.")
- Response to pain-points
- Multi-channel consistency.

## 4.    CHALLENGES FOR MULTILINGUAL VUI DESIGN

For the second part of this chapter, we turn to more design-related issues pertaining to multilingual IVRs, including the language offering, the voice of the IVR, and the input modality. Our assumption in this section is that a company or organisation has already made a decision to implement a multilingual IVR, taking cognisance of all prerequisites such as development, testing and tuning costs, maintenance and governance, and back-end integration (including integration in the call centre). Now some design choices have to be made. Our aim is to investigate what companies in South Africa are doing in this regard, and to then reflect briefly on some of these options. The aim is neither to provide exhaustive discussions (due to scope limitations), nor to provide solutions with empirical evidence (which is left for future usability research). Instead, we have chosen to list issues that one should take cognisance of (as a designer), or that could serve as stimulation for future research projects.

### 4.1. Language Offering

The language offering in a multilingual IVR is a complex issue, and is discussed elsewhere (Van Huyssteen, Sharma Grover & Calteaux, submitted) in more detail. For our current research we focussed on how many languages are offered in the investigated IVRs, what those languages are, and what strategies are being used to present the language choice. Based on our investigation, we categorised the choices that were made in the IVRs we investigated in the following categories:

- Cost driven: monolingual (English only)
- Historically driven: bilingual (English and Afrikaans)
- Linguistically driven: four languages (English, Afrikaans, isiZulu and Sesotho) – this option chooses for maximal coverage of as many speakers as possible, while still keeping costs relatively at bay; this approach is also partially in support of the National Language Policy Framework (DAC 2003).
- Demographically driven: five or six languages (additionally isiXhosa in one IVR, and isiXhosa and Xitsonga in another) – we assume that these choices were based on caller demographics
- Externally driven: eleven languages – the choice to include all official languages is driven either by political reasons or regulatory requirements.

Regarding strategies for presenting the language choice in the IVR, we found that the majority of the multilingual IVRs we investigated offered the language choice upfront, directly after the welcoming message (that was always presented in English). In two cases – both from the telecommunications domain – the language choice was offered upfront on the first call, and subsequently handled by calling line identification (CLI; i.e. the IVR associated the initial language choice with the telephone number). In one of these IVRs, the opportunity to change your language was offered in the main menu (as option 6), while the other IVR did not offer the choice again (if you want to change your language preference, you have to speak to an agent).

### 4.2. Persona Design

Persona design covers many aspects relating to the "*character*" of the voice used in the IVR system, including age, gender, accent, prosody, nature of the application, brand and corporate image, etc. We investigate two aspects of the voices selected for the IVRs we researched: accent and gender.

In the IVRs we tested, there was clear preference for a female persona over a male persona, with 25 of the 34 IVRs employing a female voice and only nine employing a male voice. The female voice was preferred across all of the domains investigated, except for the telecoms domain that preferred a male voice (see Figure 1). The

entertainment domain was neutral regarding this choice, with two of the IVRs employing a male persona and the other two a female persona.

**Figure 1:** *Gender of persona per domain*
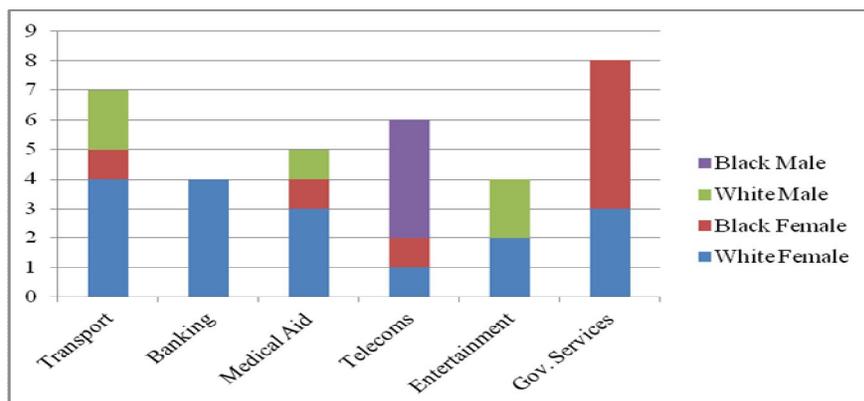


In terms of accent[2], it was found that all of the IVRs employed a South African English (SAE) accent, except for one voice in the banking sector which had a (slight) British English accent.

Within the (multilingual) South African environment VUI designers typically need to consider the following issues with regard to persona design:

- Whether to use one, multilingual voice artist, or several voice artists?
- How to deal with accented speech (e.g. is there acceptance/tolerance for Black South African English accents in commercial IVRs)?
- Whether and how to ensure consistency across personas, keeping cultural differences in mind[3]?
- Whether the persona should be more human-like and less machine-like? (See Balentine 2007 regarding this issue in western contexts.)

The choice of one or more voice artists will usually be determined by the number of languages in which the IVR is provided. It is unlikely that one voice artist would be able to speak all of the official languages with a "standard" accent. Grouping the languages together, as mentioned in 4.1 above (see caller demographics, linguistics and external drivers), might offer the option of minimising the number of voice artists used, however, attitudes towards accented speech would need to be considered (see also Cohen *et al.* 2004). The South African VUI designer therefore needs to weigh the benefits and the drawbacks of having the same voice for all or a subset of the IVRs being developed.

To our knowledge, very little – if any – South African research exists (Ndwe 2011) on language attitudes in commercial contexts (e.g. attitudes towards using one's own language, attitudes towards accented speech, specifically BSAE, etc.). The PanSALB survey (PanSALB 2000) on language use and language interaction investigated some aspects of language attitudes in commercial contexts, including language preferred when interacting with a banking institution, or at a supermarket or shop. However, the researchers who undertook the survey could not identify definite patterns with regard to language attitudes in these domains, and as a result, proposed further investigation of this aspect.

The IVRs we investigated do not appear to have mapped the persona design to the IVR's caller demographics (language, accent or culture), opting in most cases for one language, a neutral accent and a (professional) female voice. Despite not having factual data to explain the choice of persona for the various IVRs, we assume that a female voice may have been chosen in the majority of the designs in an attempt to adhere to the "likeness principle" (i.e. the principle that end-users will more probably be attracted to a persona that is similar to themselves; Nass & Brave 2005:67). In this regard, Respondent 1 remarked that female voices are often perceived by focus groups to be more neutral, i.e. more likely to be accepted across a wider range of callers from differing cultural backgrounds.

Respondent 2 noted that his company has changed from a female to a male voice in their IVR, and although they received some negative comments right after the change, acceptance now seems to be high. This suggests that the choice of gender of the persona might not be as vital as some business owners or marketing managers might perceive. Support for the latter was found from Respondent 3; for their IVR they simply employed the call centre manager's voice, which was recorded in-house, using technology available on the IVR, without consideration given to persona design or caller demographics. The main considerations were to save costs, while still getting a somewhat "*professional-sounding voice*". Other challenges, such as call centre management and business process management, seem to far outweigh concerns over persona design.

Our conclusion here is that the gender choice for a persona is probably not a critical one, especially in light of the fact that there is limited information available on language attitudes, cultural differences and gender in commercial environments in South Africa.

## 4.3. Input Modality

The debate on the choice of input modality in terms of speech input vs. touch-tone is a perennial topic in the design of VUIs. Much research (Lee & Lai 2005; Delogu *et al.* 1998; Franzke *et al.* 1993; Suhm *et al.* 2002) has been conducted in this area; Lee & Lai (2005) found that touch-tone is more suitable for simpler, linear tasks with a defined set of options, and speech input was more suited towards the more open domain and non-linear tasks which may have numerous options.

However, it was also noticed that user preference (towards speech) did not correlate with user performance (better with touch-tone) in terms of linear tasks (Lee & Lai 2005). Similar results have been observed by Delogu *et al.* (1998), Franzke *et al.* (1993), and Suhm *et al.* (2002), where the results indicate that the use of natural language (NL), which uses open-ended or free-form dialog, or at least some form of speech recognition, is favoured over touch-tone.

A point to note is that all of these comparison studies targeted business-typical applications, such as voicemail, call routing and banking, which are more applicable to users on the higher end of the technological-literacy spectrum in the developed world. Contrastingly, in an informal poll of over a thousand users of real-world information-access systems, almost half of the users responded that they would prefer to use a speech input modality "as little as possible" and only 8% would do so "most of the time" (Pearce & Bergelson 2008).

Similar comparison studies in the developing world have produced varying results, with some (Sharma Grover *et al.* 2009; Patel *et al.* 2009) reporting user preference for touch-tone, and others for speech (Sherwani *et al.* 2009). As observed by Barnard *et al.* (2008), in the developing world the user preference for a particular modality stems from an interplay of factors such as user ability (literacy, technology experience, user training) and application complexity (linear vs. non-linear tasks, open vs. closed domain).

In the IVRs we investigated, we find that a vast majority (29 out of 34) use touch-tone only (Figure 2). Of the remaining 5, only 2 are speech only with the other 3 featuring touch-tone as either a choice, fall-back or in a mixed input fashion. Strikingly, we notice that all 5 IVRs that use speech input are only available in English, thus from the IVRs examined that have multiple languages, none of them use speech input.

**Figure 2:** *Input modality used in IVRs*
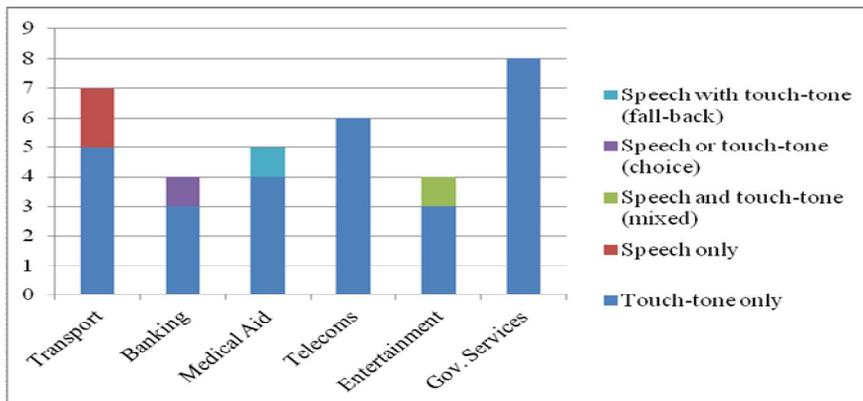


We conjecture that there are several reasons that contribute to speech input not being used in South African multilingual IVRs:

- **Availability of commercial grade recognisers** (or language resources): To our knowledge, although there has been much progress in ASR technologies being developed for the eleven South African languages (e.g. Kamper & Niesler 2011; Mbogho & Katz 2010), commercial grade speech recognisers are only available for South African English and Afrikaans, and no such recognisers are available for the remaining nine official languages (Sharma Grover *et al.* 2011c). ASR development for the African languages is majorly slowed down by the limited number of language resources (such as language models, acoustic models, grammars and pronunciation dictionaries) available for these languages. In that context, South African English and Afrikaans have leveraged on prior work done around the world on English and Dutch respectively (Sharma Grover *et al.*, 2011c).
- **Cost of commercial grade recognisers** (or language resources): In relation to the availability of technologies is also the issue of the cost of using a commercial grade recogniser, which would currently typically be built in conjunction with an international speech vendor's technology, pushing the IVR development costs even higher. As highlighted in section 3, cost seems to be the biggest driver in developing multilingual IVRs in South Africa, all the more pushing companies towards touch-tone-based solutions, which are simpler, faster and more stable to implement. Cost implications are not only limited to development and/or licensing costs, but also to maintenance costs. As Respondent 2 pointed out, if his company would decide to implement speech, the maintenance cost would probably be enormous to regularly fine-tune acoustic models, language models, grammars, etc. for multiple languages. He also pointed out that his

301

company has concerns about the availability of local expertise to contract for such maintenance work.

- **Language usage**: South Africa's multicultural environment results in South African languages being spoken with several different accents. This of course increases the time, language resources and costs involved in developing a speech recogniser that caters for the various accents for a specific South African language. This is further exacerbated by the fact that code-switching (e.g. a mix between English and an African language or between two closely-related African languages) occurs frequently in South African users' speech (Roux *et al.* 2000; Finlayson & Slabbert 1997; Slabbert & Finlayson 2002). This requires a further investment in developing appropriate speech recognisers (e.g. non-native speech recognition with multiple language models), and modules, such as language identification modules, to distinguish between closely-related languages in order to address this phenomenon.
- **End-user perception**: In the developing world, user perceptions of and inexperience with ASR technologies could also rule in favour of touch-tone over speech. For example, in some previous studies (Sharma Grover *et al.* 2009) it was found that some users considered the IVR system's voice to be a real human on the other end and not automated. Whilst using touch-tone may not completely eliminate this problem, it will certainly alleviate it as there is less human-like interaction with key presses, than with speech input.

Although one might conclude from the above that emerging markets should be thankful for escaping the so-called Jetsonian culture (see International Marketing Council of South Africa 2010), where speech is often seen as the begin-all and end-all of human-computer interaction, it is not as simple as that. The advent of smart phones and phones with touch screens are likely to have a significant impact on emerging markets, and one can therefore not simply ignore ASR as a possible input modality.

## 5.    CONCLUSION

In this research we aimed to get a better understanding of business and design issues related to IVRs in a multilingual, emerging market such as South Africa. Of the South African 34 IVRs investigated in this study only 9 had a multilingual offering (more than 1 language) and none of these multilingual IVRs had speech input. Speech input was available in 5 IVRs and only in English with the remaining 29 using touchtone as the input modality. In investigating business drivers for multilingual IVRs, we found that cost is probably the biggest driver for not implementing multilingual IVRs, over-shadowing many of the positive business

drivers, and the local availability of technologies and expertise are also significant hurdles for implementing ASR. Based on our limited interviews, our estimation is that South Africa will not see a dramatic increase in the number of commercial multilingual IVRs in the near future, unless (a) the customer demand increases; or (b) punitive laws and regulations force companies to have multilingual IVRs. Perhaps the example set by government bodies in implementing multilingual IVRs could in future contribute to strengthening the business case for commercial multilingual IVRs, while also generating valuable expertise on how to properly design and manage such IVRs.

Persona and gender choice for prompts seem to feature as low design priorities in the IVRs investigated. This could be attributed to the fact that there is a huge lack of business intelligence (or research findings) related to language attitudes (including attitudes towards cultural issues such as choice of persona), which makes it difficult for business managers and designers in industry to make informed choices during the design phase. This supports our contention that cost is a primary driver for multilingual IVR development in emerging markets. However, we also conclude that there are still huge opportunities for sociolinguistic research into the implications of language preferences, language attitudes and cultural differences for the development of (commercial and non-commercial) IVRs in South Africa.

In addition to questions for further research mentioned throughout the study, the following, amongst others, could also be addressed in future research – all within the context of commercial environments:

- **Human-computer interaction**: Do people from different cultures have different propensities towards using more machine-like or more human-like technologies? How does language choice influence metrics (e.g. call completion rates, time in IVR, etc.)? Should different metrics be applied to different user groups (e.g. groups with higher technology literacy than other groups)? Do people from different cultures or backgrounds have different levels of tolerance for queuing times? What are people's expectations or attitudes relating to music while waiting in a queue? Are customers content to switch between different languages (e.g. self-service vs. call centre agent)?

- **Language attitudes**: Do people need or prefer IVRs in their own languages, and if so, what are their expectations and requirements? What are their attitudes towards companies who do not offer multilingual services? How are IVRs that only offer a limited number of languages perceived? Will people use a closely-related language (e.g. isiZulu) if their own mother-tongue (e.g. isiXhosa) is not offered in an IVR? What are their attitudes towards accented speech (not only accented English)?

- **Scripting**: Are there significant differences between different language groups regarding discourse management in technology (e.g. turn-taking, greetings, apologising, error recovery strategies, time-outs and delays, etc.)? What challenge does tapering of prompts hold for different languages? What core terms are needed for broad application in IVRs (e.g. *hash*, *star*, *press*, *enter*, *zero*, *repeat*, *help*, etc.)? What are the most important issues related to style and register in IVR environments for different languages? How should product names be handled (e.g. translated or not)?

Issues such as these should of course not only be explored in the South African context, but also in other similar multilingual, emerging markets.

## ENDNOTES

1. See *www.internetworldstats.com/af/za.htm* and *www.google.com/publicdata*

2. Perceptions of the accent of a voice artist differ from person to person. For the purposes of this research, if at least two of the authors perceived the accent as the same (e.g. as Black South African English), then it was classified as such.

3. For example, in Germany BMW had to do a product recall for their in-car navigation system that had a female voice, as German male drivers (the vast majority of BMW drivers) did not like taking directions from females (Nass & Brave 2005).

## REFERENCES

Balentine, B. 2007. *It's Better to Be a Good Machine Than a Bad Person: Speech Recognition and Other Exotic User Interfaces at the Twilight of the Jetsonian Age.* Annapolis: ICMI Press.

Barnard, E., M. Davel & G.B. van Huyssteen. 2010. Speech Technology for Information Access: a South African Case Study. *Proceedings of the AAAI Spring Symposium on Artificial Intelligence for Development (AI-D),* Palo Alto, California.8-13.

Barnard, E., M. Plauché & M. Davel. 2008. The utility of spoken dialog systems. *Proceedings of the IEEE Workshop on Spoken Language Technology (SLT),* Goa, India.13–16.
http://www.meraka.org.za/lwazi/publications/barnard08spokendialogue

Calteaux, K., A. Sharma Grover & G.B. van Huyssteen. 2012. Business drivers and design choices for multilingual IVRs: A government service delivery case study. *Proceedings of the 3rd International Workshop on Spoken Languages Technologies for Under-resourced Languages (SLTU).* Cape Town, South Africa, May 2012.

Cohen, M.H., J.P. Giangola & J. Balogh. 2004. *Voice User Interface Design.* Boston: Addison-Wesley.

Department of Arts and Culture (DAC). 2003. *Implementation plan: National Language Policy Framework.* Unpublished. Pretoria: National Language Service.

Delogu, C., A.D. Carlo, P. Rotundi & D. Sartori. 1998. Usability evaluation of IVR systems with DTMF and ASR. *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP '98),* Sydney, Australia, November 1998.

Finlayson, R. & S. Slabbert. 1997. 'I'll meet you halfway with language': Code-switching within a South African urban context. *Languages Choices. Conditions, Constraints and Consequences,* edited by M. Putz. Amsterdam: John Benjamins.381-421.

Franzke, M., A.N. Marx, T.L. Roberts & G.E. Engelbeck. 1993. Is Speech Recognition Usable? An Exploration of the Usability of a Speech-Based Voice Mail Interface. *ACM SIGCHI Bulletin* 25(3):49-51.

Gapminder. 2007/2008. Data in Gapminder World.
http://www.gapminder.org/data/ Accessed: 30-03-2011.

International Marketing Council of South Africa. 2010. 2010 Human Development Index ratings. http://www.imc.org.za/press-room/567-2010-human-development-index-rankings.html Accessed: 30-03-2011.

Internet World Stats. 2012. South Africa Internet Usage and Marketing Report. http://www.internetworldstats.com/af/za.htm Accessed: 30-03-2011.

IT News Africa. 2009. Africa's high mobile penetration sets the stage for Internet revolution. http://www.itnewsafrica.com/2009/07/africas-high-mobile-penetration-sets-the-stage-for-internet-revolution/ Accessed: 30-03-2011.

Kamper, H. & T. Niesler. 2011. Multi-accent speech recognition of Afrikaans, Black and White varieties of South African English. *Interspeech 2011: Proceedings of 12th Annual Conference of the International Speech Communication Association.* Florence, Italy. 28-31 August 2011.

Lee, K.M. & J. Lai. 2005. Speech Versus Touch: A Comparative Study of the Use of Speech and DTMF Keypad for Navigation. *International Journal of Human–Computer Interaction* 9(3):343-360.

Mbogho, A. & M. Katz. 2010. The impact of accents on automatic recognition of South African English speech: a preliminary investigation. *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT)*, Bela Bela, South Africa.187-192.

Nasfors, P. 2007. *Efficient voice information services for developing countries.* Master's thesis, Department of Information Technology. Uppsala: Uppsala University.

Nass, C. & S. Brave. 2005. *Wired for Speech: How Voice Activates and Advances the Human-Computer Interaction Relationship.* MIT: Cambridge, USA.

Ndwe, T. 2011. *Usability Engineering of Interactive Voice Response (IVR) Systems in Oral Users of Southern Africa.* PhD Thesis. Cape Town: University of Cape Town.

PanSALB. 2000. *Language use and language interaction in South Africa – A national sociolinguistic report.* Unpublished research report. Pretoria: Pan South African Language Board.

Patel, N., S. Agarwal, N. Rajput, A. Nanavati, P. Dave & T.S. Parikh. 2009. A comparative study of speech and dialed input voice interfaces in rural India. *Proceedings of the Conference on Human Factors in Computing Systems (CHI'09)*, Boston.51-54.

Pearce, T. & M. Bergelson. 2008. Alignment index for speech self-service. Dimension Data Technical Report. http://www.dimensiondata.com/NR/rdonlyres/9191A848-5F35-459F-8239-8D9D2248414E/8791/mainstreamspeechalignmentindexreport2 Accessed: 15-09-2008.

Plauché, M. & U. Nallasamy. 2007. Speech Interfaces for Equitable Access to Information Technology. *Information Technologies and International Development (ITID) Journal* 4(1):69-86.

Roux, J.C., E.C. Botha & J.A. du Preez. 2000. Developing a Multilingual Telephone based Information System in African Languages. *Proceedings of the Second International Language Resources and Evaluation Conference.* Athens, Greece, 31 May – 2 June 2000. http://www.ast.sun.ac.za/publications/lrec2000-jcrebjadp.pdf

Sharma Grover, A., M. Plauché, C. Kuun, & E. Barnard. 2009. HIV health information access using spoken dialogue systems: Touchtone vs. Speech. *Proceedings of the International Conference on Information and Communications Technologies and Development (IEEE),* Doha, Qatar.95-107.

Sharma Grover, A., O. Stewart & D. Lubensky. 2009. Designing interactive voice response (IVR) interfaces: Localisation for low literacy users. *Proceedings of the 12th IASTED International Conference on Computers and Advanced Technology in Education (CATE),* St. Thomas, Virgin Islands.673-680.

Sharma Grover, A. & E. Barnard. 2011a. Comparing Two Developmental Applications of Speech Technology. *Conference on Human Language Technology for Development 2011.* Alexandria, Egypt.81-86.

Sharma Grover, A. & E. Barnard. 2011b. The Lwazi Community Communication Service: Design and Piloting of a Voice-based Information Service. *Proceedings of the 20th WWW Conference,* Hyderabad, India.433-442.

Sharma Grover, A., G.B. van Huyssteen & M.W. Pretorius. 2011c. The South African Human Language Technology Audit. *Language Resources and Evaluation* 45(3):271-288.

Sherwani, J., S. Palijo, S. Mirza, T. Ahmed, N. Ali & R. Rosenfeld. 2009. Speech vs. touch-tone: Telephony interfaces for information access by low literate users. *Proceedings of the International Conference on Information and Communications Technologies and Development (IEEE),* Doha, Qatar.447–457.

Slabbert, S. & R. Finlayson. 2002. Code-switching in South African townships. *Language in South Africa* edited by R. Mesthrie. Cambridge: Cambridge University Press.235-257.

Suhm, B., J. Bers, D. McCarthy, B. Freeman, D. Getty, K. Godfrey, & P. Peterson. 2002. A Comparative Study of Speech in the Call Center: Natural Language Call Routing vs. Touch-Tone Menus. *Proceedings of the Conference on Human Factors in Computing Systems (CHI'02)*, Minneapolis, Minnesota, USA.283-290.

Tucker, R. & K. Shalonova. 2004. The Local Language Speech Technology Initiative – localisation of TTS for voice access to information. *Proceedings of the SCALLA Conference,* Kathmandu, Nepal, January 2004. http://www.elda.org/en/proj/scalla/SCALLA2004/tucker.pdf

United Nations Conference on Trade and Development. 2010. South-South cooperation offers new opportunities for transforming African economies. http://www.unctad.org/Templates/StartPage.asp?intItemID=2871&lang=1 Accessed: 30-03-2011.

Van Huyssteen, G.B., A. Sharma Grover & K. Calteaux, submitted. Offering multiple languages in interactive voice response systems.