Employing Apple's Siri to practice pronunciation: A preliminary study on Arabic speakers*

*Martin Molden***
The New York Public Library, USA

Abstract
This paper reports on a preliminary study on Arabic Speakers using Siri® to practice pronunciation and aims to uncover Siri's positive affordances for language learning in relation to its corrective feedback and interpretation of speech. The implications of this study can be relevant to researchers on Computer-Assisted Language Learning, classroom teachers, speech recognition software developers, and language learners.

Introduction

Speech Interpretation and Recognition Interface, Siri®, is an artificial intelligent speech recognition software developed by Apple Inc. It is pre-loaded on any Apple device succeeding iPhone 5 and serves as a personal assistant. Some of its functions include setting the alarm, giving directions, recognizing music, defining words, determining calorie content in foods, locating contacts, reading and sending texts or emails, and pointing out the closest restaurants, Wi-Fi hotspots, or local businesses assorted by rating (Moore, 2014). Although Siri was developed in order to equip Apple's customers with a personal assistant, it can also be used for the purpose to practice pronunciation. In order to carry out its users' requests, it is critical for Siri to understand its interlocutor's utterances. Thus, Siri can be used to teach almost any oral language feature with respect to intelligibility through negotiation for meaning (but limited by what languages are available through Apple).

When it comes to pronunciation intelligibility, it is important to keep in mind Celce-Murcia, Brinton, and Goodwin's (2010) point that even if a speaker incorporates perfect vocabulary and grammar usage, a certain pronunciation threshold has to be reached in order to avoid communication breakdowns. At the same time, Jenkins (2000) pointed out how mutual intelligibility, not perfect pronunciation, should be the ultimate goal of language teaching in general, given that English has become a global phenomenon and the number of speakers of English as a Lingua Franca has far exceeded the number of speakers of English as a native language.

Figure 1 illustrates how Siri relates to Jenkins's statement. First, speech is evaluated on whether it is comprehensible or not. If Siri understands what the speaker is trying to say, it will carry out the requested function. However, if Siri does not understand, it will either execute another function, state it does not understand, or search the web for answers. This means that comprehensible output does not always equal perfect pronunciation. Siri is context-aware, just as human interlocutors are, and knows how to interpret small errors that do not interfere greatly with meaning.

Molden, M. (2015). Employing Apple's Siri to practice pronunciation: A preliminary study on Arabic speakers. *Hawaii Pacific University TESOL Working Paper Series 13*, 2-17.
Website: http://www.hpu.edu.
* This paper is based on a presentation given at Hawaii TESOL in February, 2015.
** Email: mmolden@my.hpu.edu. Address: TESOL Program, MP 441, 1188 Fort Street Mall, Honolulu, HI 96813, USA.

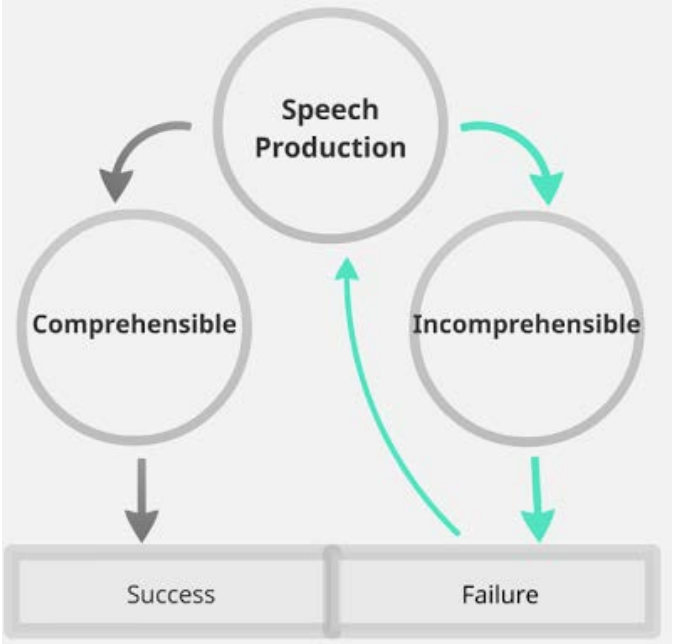

Figure 1. Model of Siri's intelligibility pattern and how it involves negotiation of meaning

This paper reports on a small-scale study of how two Arabic learners of English as a second language used Siri to practice English pronunciation. Based on my observations, I discuss Siri's positive and negative affordances in relation to its accessibility; its base in intelligibility, its context dependency, its feedback, its validity, and its negotiation for meaning. Before I proceed, it is necessary to revisit the notion of corrective feedback in second language learning, since this is a key feature of Siri that renders it useful for pronunciation practice.

Corrective Feedback with Speech Recognition Software

According to Eskenazi (1999), corrective feedback occurs in two instances: in natural conversation where the speakers' relationship allows them to correct each other's errors or when the intended meaning does not get across. In both of these cases, negotiation for meaning is a necessary process. The way Siri does the former is by transforming its interlocutor's speech into written discourse. If the request is spelled correctly, it is likely to also have been pronounced correctly. The latter, on the other hand, is accomplished by analyzing Siri's reaction. If it carries out the request, the pronunciation was intelligible, and if it does not, the pronunciation was probably faulty. Further, according to Kawai and Hirose (2000), "Every pronunciation CALL system should explain to the learner (a) what his mistake was, (b) the severity of the error, and (c) how to correct his mistake" (p. 142). Even though Siri does not seem to explain any of these explicitly, it does implicitly point out the mistake (by means of spelling) and the severity of the error (by means of intelligibility). It does not, however, explain to the learner how to correct these mistakes.

Kawai and Hirose (2000) conducted research on a speech recognition software that was capable of giving such feedback. The software was developed in order to assist teaching Japanese as a second language teachers in their pronunciation teaching on a phoneme level. Kawai and Hirose decided to focus on the duration of Japanese double-mora phonemes in contrast with single-mora phonemes, due to them being, vowel duration excluded, spectrally identical and particularly troublesome for

Japanese language learners to acquire and retain. In order to observe the software in action, the researchers had the participants read minimal pairs containing the target vowel duration. The software was based on experiments on native Japanese speakers' perception of confusability in relation to these vowel durations and may on this basis provide students with intelligibility scores, instructions on how to improve pronunciation, and their vowel duration in milliseconds. Kawai and Hirose's (2000) study concluded that the "learners quickly capture the relevant duration cues," but that does not necessarily mean that they will retain these in later, more informal settings.

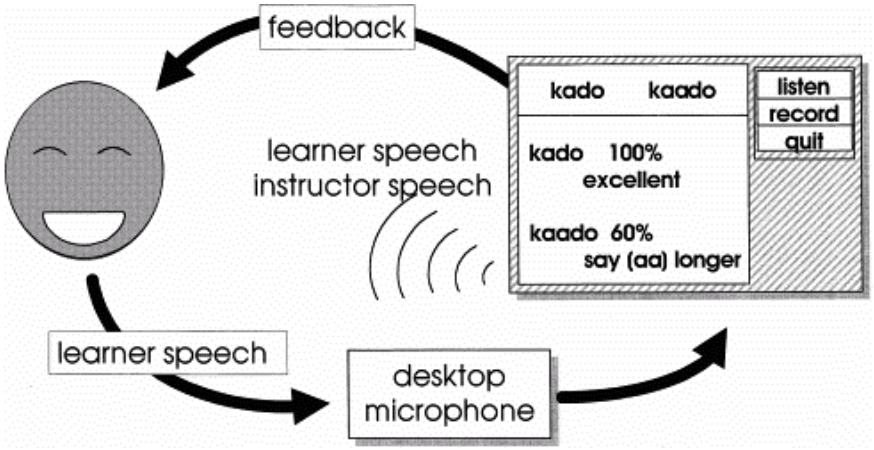

Figure 2. System-user interaction in Kawai and Hirose's (2000) study (p. 134)

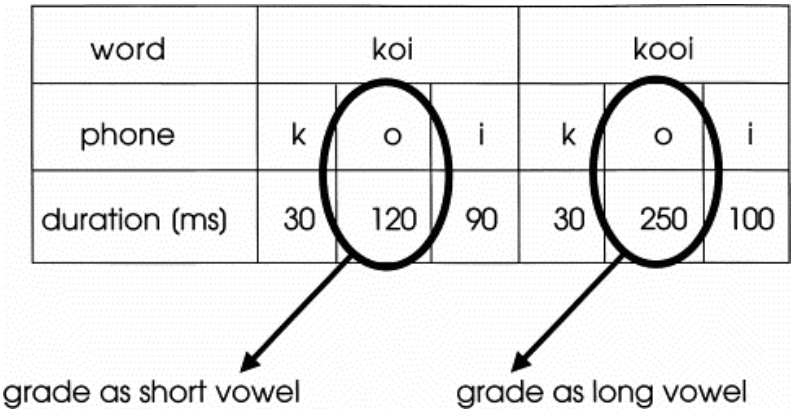

Figure 3. Duration measurement and evaluation in Kawai and Hirose's (2000) study (p. 134)

Figure 2 illustrates how Kawai and Hirose's (2000) software models the target phoneme, interprets its users' speech production, and provides corrective feedback. Figure 3, on the other hand, breaks down the foundation on which the software's duration measurement was built. For instance, the phoneme pair *koi* and *kooi* is distinguished by how many milliseconds the *o*-sound is produced. If it was pronounced for 120 milliseconds, it would result in a perfect pronunciation of *koi*; if it was pronounced for 250 milliseconds, it would result in a perfect pronunciation of *kooi*. If the duration of the o-sound falls anywhere else along this continuum of milliseconds, the software will provide

feedback on how to get closer to the phoneme duration benchmark in point. These benchmarks – 120 and 250 milliseconds – came to be as a result of testing native Japanese speakers' perception of speech samples of varying phoneme durations.

Dalby and Kewley-Port's (1999) and Kawai and Hirose's (2000) studies gave much appreciated insight into how these speech recognition software work. However, while decontextualized minimal pair drills may have served its purpose for research purposes, it may not be as efficient in language learning. The attention that is drawn to the target phoneme may indeed help the student produce it, but in naturally occurring speech, this attention is spent on content and negotiating of meaning as well as sound production. Therefore, in order to develop phoneme pronunciation skills that are applicable to the outside world, research has to be done on how language learners can negotiate for meaning using speech recognition software in a contextualized and communicative manner.

Apple's Siri may be a suitable tool for pronunciation in tasks that are similar to real-life communication because, as Boyle (2013) stated, Siri is both context aware and interprets and corrects small grammar errors automatically. This feedback, however, is of less quality than both Kawai and Hirose's (2000) and Dalby and Kewley-Port's (1999) software, but Siri is also much more accessible and familiar to most students. It is also more likely to be used outside of class due to its many functions that do not necessarily focus on language features explicitly. Pronunciation becomes a means to an end, targeting intelligibility and comprehensibility.

Research Questions

In this preliminary study, I aim to find out the answers to the following questions:

1. What forms of corrective feedback does Siri provide?
2. To what extent were communication breakdowns between the learners and Siri due to the learners' pronunciation errors?
3. What do learners like and dislike about using Siri for pronunciation practice?

Unlike Kawai and Hirose's (2000) and Dalby and Kewley-Port's (1999) studies, which focused on minimal pairs, the present study aims to use Siri for pronunciation practice in context through the use of meaningful commands.

Methodology

Participants

In order to ensure reliable testing, and with an interest in phoneme difficulties based on language interference, two participants with the same language background were selected. Participant #1 (P1) was a 26-year-old man with an upper-intermediate command of English. He was born and raised in Riyadh, Saudi Arabia, but was residing in Honolulu, Hawaii at the time of the study. He was relatively familiar with technology in general, and he had an iPhone 5s on which he used Siri on a weekly basis to determine the weather conditions, type sizable texts, or spell words he did not know how to spell. Participant #2 (P2) was a 24-year-old woman with a low intermediate command of English. She was born and raised in Riyadh, Saudi Arabia, but was residing in Honolulu, Hawaii at the time of the study. She did not speak English until she moved to Hawaii two months prior to this study, but she was relatively familiar with technology in general and had an iPhone 6 on which she used Siri approximately three times a week to search Google or call her contacts.

Instruments and Data Collection

In this study, an instrument with two parts (see Appendix A) was used to elicit the participants' pronunciation. The first part consisted of guided practice activities centered on a set of segmentals

that Arabic speakers were likely to find challenging. These were determined based on a contrastive analysis of the Arabic and the North American English phoneme inventories as well as phonotactics. After narrowing these difficulties down, /ɪ/ and /ɛ/ stood out as the most difficult phoneme pair, due to the absence of /ɛ/ in Arabic vowel inventory. The contrastive analysis also uncovered potential difficulties with consonant clusters due to Arabic not having consonant clusters in initial position or more than two consecutive consonants in any other position. /ɛ/ could therefore be replaced by /ɪ/ in words like *melon*, and consonant clusters would probably be characterized by a process of vowel sound insertion called anaptyxis. In the tasks given to the participants, I made sure that not all questions and requests contain the targeted problematic segmentals. This was done deliberately in order to both lower the participants' affective filters somewhat by making sure they could execute at least some tasks and also in order to make the purpose of the study less obvious. I was mindful that since the tasks only required the students to read sentences aloud, English orthography could lead to difficulties. This is because the Arabic orthography is close to phonemic (its graphemes are closely related to the phonemes of the language) but this is not the case with English. English tends to have graphemes represent several phonemes that could overlap with other graphemes, while also presenting graphemes that are pronounced in some settings, but not in others. Many Arabic Speakers of English (ASEs) could therefore, when they encounter new words, pronounce every grapheme they see when they read.

The second part of the data collection consisted of free practice activities based on both a principle of making learning authentic and the following statement made by Boyle (2013):

> … you could let them discover the natural way to ask the question on their own, through trial and error. This sort of experimentation is a vital and necessary part of the learning process, but one that obviously involves lots of mistakes. Mistakes made in class are often traumatizing to students and this only slows or even stops the learning process. Siri offers a way to turn mistakes into a form of play.

In other words, the second part was designed in order to elicit the participants' pronunciation of specific sounds in context, embedded in meaningful activities.

An American female voice was also chosen for both of these parts in order to make the testing coherent and cater to the English standard the participants were most frequently exposed to. Each part took the participants approximately ten minutes to complete.

Additionally, the participants were interviewed using the questions in Appendix B in order to find out their (a) subjective likes and dislikes toward Siri, (b) understanding of the communication breakdowns during the tasks, and (c) suggestions for future language teaching using Siri.

Analytical Procedure

As an assessment of the technology's validity, a communication breakdown chart was utilized to determine whether these breakdowns happened due to pronunciation deficits or technology limitations (See Appendix C). Furthermore, this analysis considered such communication breakdowns to be instances where the interlocutor's speech production did not lead to an execution of the intended function; thus, pronunciation errors that did not lead to a misinterpreted or otherwise failed execution of the function in point were not tallied. Neither did breakdowns that occurred due to the participants' Siri missing crucial information required for carrying out certain tasks; these could include knowing who the owner of the phone was in order to create a nickname or establish relations to other contacts in the phone's contact list, or establishing a working email in order to send mails. Further, identical consecutive breakdowns as a result of a participant trying to negotiate breakdowns counted as only

one breakdown. I used my training in linguistics and language teaching to determine whether a communication breakdown was due to a pronunciation deficit or technology limitation. If an utterance contains pronunciation that differ from standard English, I logged the breakdown as being due to pronunciation deficit. On the other hand, if an utterance sounds native-like but there was a communication breakdown, I categorized the cause as technology limitations. Finally, those errors that could be accredited both pronunciation deficits and technology limitations would only count as one breakdown and classified as either one of these depending on which one was the most prominent.

Findings

Forms of Feedback Provided by Siri

Siri does not provide feedback when the learner mispronounces a sound but there is enough intelligibility in context. An example of how Siri provides corrective feedback is when P1 asked Siri if it could tell him a joke. A native speaker would pronounce the phrase as /tɛl miy ə dʒowk/ while P1 produced /tɪl miy ɑ dʒɔk/. P1 had a slight pronunciation error here: as was expected from the contrastive analysis between English and Arabic (see above), P1 had problems differentiating /ɛ/ and /ɪ/, in addition to not producing the off-glide after /ɔ/. However, he also provided Siri with sufficient context for it to understand what he was trying to say. Indeed, Siri understood what he was trying to say, judging from the context, and carried out P1's request by telling P1 a joke.

In contrast, in cases where less context is provided, Siri seems to be less lenient with pronunciation deficits, just like human listeners would. The kind of feedback it provides is implicit, in the form of a misunderstanding of the learner's request. Example 1 shows how a more isolated word can lead to confusion. In this task, the participants were prompted to ask for a definition of the word *bet*.

Example 1

P1: /wə t das bɪt miyn/
Siri's response: From 17 definitions of "big"…

Siri also provides feedback through transcription of what the speaker is saying. If a word is pronounced incorrectly, it was often spelled incorrectly, as in Examples 2-4.

Example 2

Original prompt: How many ounces are there in one cup?
P1: /haw maniy awnə sə s ar dɛr ə n wə n kə p/
Siri's transcription: How many all bonuses are there in one cup?

P2: /haw maniy ansə s ar dɛr ə n ankɔ/
Siri's transcription: How many ounces are there in uncle?

Example 3

P1's intended phrase: How many calories are there in a melon?
Production: /haw mɛniy kæloriyz ar dɛr ə n ə mɪlɔn/
Siri's transcription: How many calories are there in the middle?
Siri's reaction: Did not carry out the intended request

Example 4

P2's Intended phrase: Find a good restaurant nearby
Production: /faynd ə god rɛstʊrant nɪyrbey/
Siri's transcription: Find a good restaurant near me
Siri's reaction: Carried out the intended request

In these examples, Siri pointed out the errors by transcribing a word or phrase that was not intended. Notice, however, that only the words that may limit intelligibility got corrected. No native speaker would say /god/ or /rɛstʊrant/ for *good* and *restaurant*, but they are perfectly understandable in context; thus, they did not get corrected by Siri. Note also how these words are pronounced very close to their orthographic spelling, as mentioned before.

Furthermore, Siri's reaction provides feedback to the learner how severely the pronunciation deviates from its standard. Even though there seems to be a continuum in human speech perception, Siri either carries out the request or fails to do so based on its intelligibility. Siri's judgement seems to be based on context and collocation rather than the degree to with individual sounds differ from standard pronunciation. Thus, both learners in Examples 3 and 4 produced vowel deviations, but P2's request was carried out while P1's was not.

Thus, Siri seems to address only two out of the three criteria for feedback stated by Kawai and Hirose (2000) mentioned earlier: "(a) what his mistake was, (b) the severity of the error, and (c) how to correct his mistake" (p. 142). Siri addresses what the speaker's mistake was by spelling words incorrectly or inserting words the speaker had not intended; it addresses the severity of the error by carrying out the function or failing to do so, but it never seems to provide any explicit information on how speakers can improve their pronunciation. It never seems to provide explicit suggestions as to how they should proceed in improving their production. Thus, even though Siri's feedback may have some positive affordances, they have to be seen in relation to its negative affordances.

Kawai and Hirose (2000) pointed out how their speech recognition software addressed all of the criteria mentioned above, making their software's feedback superior to Siri's. However, their software falls short in language teaching in that it does not provide sufficient context for the phonemes in point, nor is it as accessible to the language learners as Siri.

As is indicated with turquoise arrows in Figure 1 above, Siri accomplishes this by having speakers modify their speech whenever it fails to carry out their requests. P1 had an instance in the second part of the test that can illustrate this aspect of Siri's positive pronunciation affordances (Example 5).

Example 5
Intended function: Having Siri read a text out loud
Attempt #1: /lɪsə n tuw e tɪkst/
Siri's response: Web search for Listen.txt (failure)
Attempt #2: /lɪsə n tuw e tɪkst/
Siri's response: Searching music for hottie (failure)
Attempt #3: /lɪsə n tuw e tɪkst/
Siri's response: Opening a Wikipedia article on dicks (failure)
Attempt #4: /lɪsə n tuw e tɛkst/
Siri's response: Reading a text out loud (success)

As was expected from the contrastive analysis of phoneme registries, P1 confuses /ɛ/ with /ɪ/ in the word text. This pronunciation deficit is something that most ASEs are unaware of and struggle

with, so there is no wonder that P1 keeps making these mistakes. After substituting /ɛ/ with /ɪ/ three times without Siri carrying out the intended function, however, P1 modified his speech slightly by replacing /ɪ/ with /ɛ/ and gets the result he wanted. Even though this negotiation of meaning could lead to improvements in pronunciation, it is neither mutual, nor always based in pronunciation deficits.

In short, Siri's feedback, then, is based on intelligibility through its context-dependency and can help improve its users' pronunciation in instances where negotiation of meaning occurs. Siri also points out the error and, to some extent, the severity of it, but fails to provide explanations on how to treat it.

Communication Breakdowns and Pronunciation Errors

As was mentioned earlier, Jenkins (2000) proposed that all language teaching should be based on mutual intelligibility. She does not address speech recognition software particularly, but her proposal is built upon having two interlocutors arrive at a common understanding. With Siri, this process is not mutual; even though speakers may change their pronunciation in order to negotiate meaning, Siri does not seem to negotiate with them. Besides, the communication breakdowns that occur do not always originate in pronunciation deficits. My analysis of P1's and P2's communication breakdowns was based on whether they were due to pronunciation deficits or technology limitations.

Based on these data, technology seems to be the cause of less than half of the communication breakdowns, but the rate can vary greatly from participant to participant. P1's breakdowns were caused by technology limitations in only 11% of the cases, compared to 43% for P2. One possible reason for this could be that P2 generally had a clearer pronunciation, especially in distinguishing /ɛ/ from /ɪ/. In any case, more extensive research is needed in order to determine why the technology interference varies this greatly.

The above analysis was made on the assumption that even native speakers of English encounter communication breakdowns when using Siri, and there are even more frequent communication breakdowns for a non-native speaker. The participants in this study, however, seemed to come across relatively few communication breakdowns caused by technology limitations, which may indicate that this is not such a prevalent issue as previously assumed.

	Total number of breakdowns	Breakdowns due to pronunciation deficits	Breakdowns due to technology limitations
Participant #1	9	8	1
Participant #2	7	4	3

Table 1
Communication Breakdowns Sorted by Cause

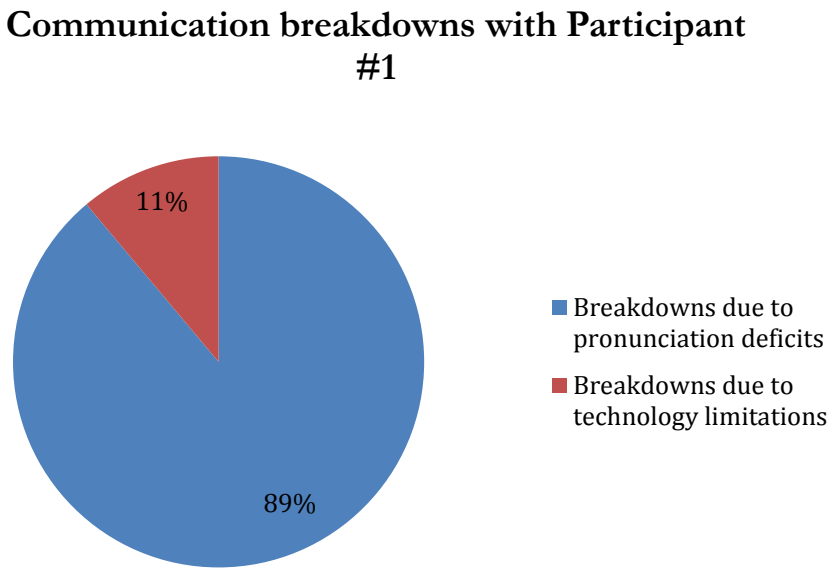

Figure 4. Causes of communication breakdowns with Participant #1

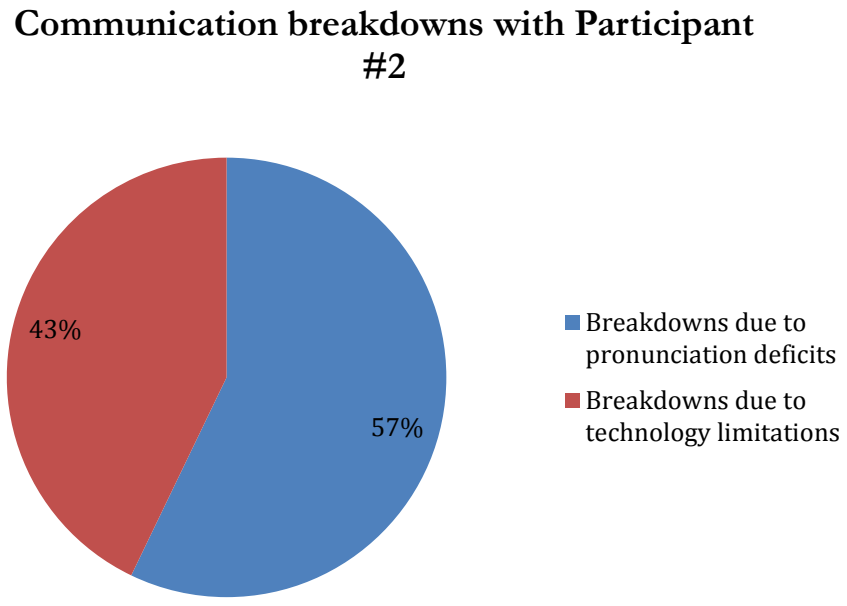

Figure 5. Causes of communication breakdowns with Participant #2

In sum, this preliminary study indicates that pronunciation deficits are the main cause of communication breakdowns when second language learners interact with Siri. More extensive research is called for in order to confirm this preliminary finding.

Learners' Perception about Siri as a Pronunciation Practice Tool

When asked whether they were familiar with Siri, both of the participants responded that they were. They both said they talked to it weekly and used it to perform a range of functions. P1 explained that he used Siri in order to determine the weather, write his texts for him, or show him the spelling of words he knew how to pronounce, but not to spell. P2, on the other hand, said she used Siri in order to search on Google or call contacts, while both of them stated that they used Siri for fun whenever they were bored. Most teenagers and young adults, like the language learners in this study, carry their smartphones around with them wherever they go, and may turn to them whenever they have a spare moment. Siri may, therefore, meet the students where they are and aid classroom pronunciation teaching outside the classroom in a way that students find fun and motivating. Regardless of the learners' purpose, Siri requires them to speak a lot while focusing on pronunciation and the meaning they are trying to convey in order to carry out these functions.

Besides, even though both of the participants in this study had capable devices and could access Siri with ease, there are those who do not have any up-to-date Apple devices. These language learners may have to resort to other software depending on their devices. Android users, for instance, may use Google Now, while Windows users may use Cortana, both of which are speech recognition software and not much different from Siri. Language learners who do not have any of these may have to pair up with somebody who does.

When the participants were asked what they liked about the software, they responded that they liked the range of functions that Siri could carry out for them, and that all of these are accessible in one application. On the other hand, they also thought Siri was a bit picky with their pronunciation and too quickly misinterpreted their requests instead of asking them what they meant to say. Both of them also pointed out that it is a shame that Siri does not speak Arabic even though Apple Insider (2014) and Cooper (2012) point out how Arabic among many other languages may be added in the not too distant future. When the participants were asked why they thought Siri sometimes misinterpreted them, they responded that it might be because they sometimes struggle with vowels, that they pronounce words like "Saudi Riyals" using an Arabic pronunciation, or that Siri is just very selective about pronunciation. Furthermore, when they were asked if they thought Siri could be used for language teaching, P1 responded that he thought Siri could be good for learning how to spell new words, while P2 thought Siri could be used to improve speaking only.

In brief, the participants thought that Siri was easily accessible and liked the number of functions it could carry out for them but disliked how picky it could be with their pronunciation and how it did not speak Arabic yet.

Discussion and Conclusion

Like any technology, Siri has its positive and negative affordances. It is popular and easily accessible, but can only be found on relatively recent Apple devices; it provides feedback based on intelligibility, but the quality of this feedback may be up for question; and, even though certain communication breakdowns seem to foster changes in pronunciation, these breakdowns may also originate from other factors than pronunciation deficits. It is important to keep in mind, however, that Siri was never intended to be used as a language learning tool; it was its need for comprehensible input that united the developers' and language learners' purpose.

With that being said, a teacher incorporating Siri in pronunciation teaching needs to consider if the students have compatible devices. If the number of students without compatible devices is minimal, these students can group up with those who do; if the number of students without

compatible devices is high, the teacher's device can be used in, for instance, station-work oriented tasks. The fact that both of the participants in this study owned modern Apple devices was fortunate but must not be taken for granted in a classroom of numerous students. Android and Google smartphones, for instance, are relatively popular as well. Luckily, if they are up-to-date, these devices also come with speech recognition software that can complement or substitute Siri whenever needed.

Furthermore, speech recognition software are developing rapidly, so teachers should also find out if there has been recent updates or improvements to the software in point. If Siri, for instance, has been equipped with an explicit form of feedback that tells the students what they need to do in order to improve their pronunciation, its feedback would be much more valuable and address all of Kawai and Hirose's (2000) criteria. If no such improvement has been made, however, Eskenazi (1999) pointed out how the teacher should complement Siri with explicit instruction on phonemes and how to articulate them.

The kind of language students use with Siri also needs to be taken into consideration when a teacher incorporates Siri, as the language demands in interaction with Siri may be quite minimal. Students may learn how to make requests and commands, but features such as articles and markers of politeness can easily be dropped without limiting intelligibility; thus, teachers may want to focus on naturally occurring speech instead of Siri to develop students' pragmatic abilities.

Even though Siri has both negative and positive affordances, it is developing rapidly, and it is also likely to appear in new places with new functions in the future. Apple is releasing Apple Watch in early 2015, and several technology websites have been speculating about Apple developing iGlasses, both of which are likely to incorporate Siri. Siri is also likely to have more features in IOS 9, but even if these were only theories and assumptions, Rogerson (2014) argued that a more humanized Siri would be probable due to Cortana already incorporating inflections. Furthermore, seeing how far the speech recognition software in Kawai and Hirose's (2000) and Eskenazi's (1999) studies have come in terms of feedback, it could only be a matter of time before this technology is accessible to language teachers.

This exploratory study is not without limitations. First, the data included only two participants. According to the Law of Large Numbers, quantitative studies could determine a more reliable and accurate percentage of overall technology interference in communication breakdowns (Bernoulli, 1713). Second, the identification of the causes of breakdowns was done by only one researcher. Finally, the tasks used were highly structured and controlled. Future research would benefit from establishing inter-rater reliability and a larger sampling of participants using freer tasks where subjects are asked to interact with Siri without any given guidelines.

References

Apple Insider. (2014, June 23). Apple looking to expand Siri localization with at least 9 new languages. *Apple Insider*. Retrieved from http://appleinsider.com/articles/14/06/23/apple-looking-to-expand-siri-localization-with-at-least-9-new-languages

Bernoulli, J. (1713). *Arts conjectandi* [Arts of conjecturing]. Basel, Switzerland: Impensis Thurnisiorum.

Boyle, M. (2013, September 23). Did Apple accidentally invent a cool language learning app? Retrieved from http://eltjam.com/did-apple-accidentally-invent-a-cool-language-learning-app/

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation - A course book and reference guide*. New York, NY: Cambridge University Press.

Cooper, C. (2012, November 20). Siri about to speak in (new) tongues? *CNET*. Retrieved from Cnet Web Site: http://www.cnet.com/news/SIRI-about-to-speak-in-new-tongues/

Dalby, J., & Kewley-Port, D. (1999). Explicit pronunciation training using automatic speech recognition technology. *CALICO Journal*, 425-445.

Eskenazi, M. (1999). Using automatic speech processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language Learning and Technology*, 62-76.

Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford, UK: Oxford University Press.

Kawai, G., & Hirose, K. (2000). Teaching the pronunciation of Japanese double-mora phonemes using speech recognition technology. *Speech Communication*, 131-143.

Moore, Q. (2014, October 1). How to use SIRI - Full list of Siri commands for iPhone, iPad, Video. *TechBlog*. Retrieved from http://techblog.tv/full-list-of-SIRI-commands-how-to-use-SIRI/

Rogerson, J. (2014, October 2). IOS9: What we want to see. *Techradar*. Retrieved from http://www.techradar.com/us/news/software/operating-systems/ios-9-what-we-want-to-see-1253732

Appendix A

Tasks with SIRI

Part 1: Say "Hey SIRI" to start. Then, ask SIRI the following questions/requests:

1. Tell me a joke
2. Do I need an umbrella today?
3. Call me Ben
4. My wife/husband is [name of somebody in your contact list]
5. What month is it?
6. How many days until Veterans Day?
7. What does "bet" mean?
8. What is 3 times 23?
9. What is the square root of 133?
10. How many ounces are there in one cup?
11. How much is $20 in Saudi Riyals?
12. What is the tip on $43?
13. How many calories are there in a melon?
14. How far away is Saudi Arabia?
15. What is the population of **Serbia**?

Part 2: Do the following using SIRI:

1. Send an email to somebody
2. Listen to a text
3. Send a text to somebody
4. Find a good restaurant nearby
5. Find out which movies are being released this month
6. Make a reminder

Appendix B

Communication Breakdown Charts

Task 1: Guided Practice		
Total # of tokens		
SUM:		
# of tokens caused by…		
Pronunciation deficits in subject	Deficits in technology	Other means
SUM:	SUM:	SUM:

Task 2: Free Practice		
Total # of tokens		
SUM:		
# of tokens caused by…		
Pronunciation deficits in subject	Deficits in technology	Other means
SUM:	SUM:	SUM:

Appendix C

Interview Questions

1. What did you like about SIRI?

2. What did you dislike about SIRI?

3. Did SIRI have difficulties understanding you at some point? What do you think the reason may be?

4. Do you think SIRI could help you learn English? If so, how?