

# Transcription Case Study

## Client:

The client Linguistix Tank utilizes an international community of on-demand professional linguists to provide real-time transcription, voice-data collection, hand writing recognition and localization services across the globe. They help companies leverage the power and complexities of language to develop better language recognition software, curate more linguistically appropriate marketing materials, and enter new markets with the confidence that their brand is effectively communicating to their target audience.

Comprised of over 1,000 professional linguists from around the world, their growing community of local language experts represents one of the most adaptive, diverse, and unique teams across today's language services industry.

## Problem:

Linguistix Tank has a running project of bettering voice language recognition for a certain search engine company. This project is dedicated to various Arab countries with different dialects. This requires a huge number of Arabs who can understand Arabic with its different dialects to be able to work on this project as transcribers.

There was a main problem which was choosing the suitable applicants from a huge number that was more than 10,000 applicants. Then there is a step of training them for the required job. The client did not have a systematic way to get this done. The applicants used to consume a huge amount of time during the recruitment and hiring procedure. It used to take the company time to filter the applicants, choose those who are suitable and those who are not, and to evaluate them accordingly. Then comes another main problem which is training. Since most of the applicants are going to work online, it is normal for the applicants to be from various locations and need to be trained online. Linguistix Tank did not have its training material hosted online, and could not figure out a way to organize and grant access to those who were going to be trained, and had to train candidates in groups via online sessions.

This led to hiring people who weren't suitable for the job, a high turnaround rate, and inefficient use of time. The final product was low quality. This negatively impacted the purpose of the project which was to train the search engine for better voice recognition.

Another problem was linking the individual productivity to the payment. At the end of each task since tracking the progress of each individual on every project was a hassle that led to time wasted on debates regarding the number of utterances transcribed.

## The Challenge:

The challenge was creating a system that should solve all the previous problems and linking this system to the proprietary data entry system the search engine company already had. This required a successful process of data migration. The required system needed to be constructed in an organized way, with levelled stages. Each stage should solve a problem the client is facing.

Another challenge was creating a platform that hosts the client's training material. Going through the training should be done individually for each applicant, in order to give individual access to the second step which is a screening test to filter out applicants.

Linking the achievement of each individual and his/her record count to the real payment was also a challenge, since it needed to be linked to the reporting system in which the transcribers will use. It also needed to be very secure and accurate.

## The Solution:

We offered a platform that the applicants can access individually, each with his/her own account. The system starts with a gateway to the training material. After each applicant goes through the training material, it is possible to reach a test. This test we prepared by extracting real audio samples from the system the applicants would be using. The system then allows for test administrators to review tested applicants and determine whether or not they should be hired. The systems allows for filtering tested candidates according to the score they achieved on the test, the duration taken to complete the test, or when they actually took the test. It also provides a powerful solution for the company to choose the most suitable candidates for the job.

Actions	Pera_pera	Notes	Score	City	Mobile	Email	Name
Test \$ ✖	bt-test-arxlevant-25 ✖	Add note	103	Cairo	01149232795	abdelrahman25196@gmail.com	Abdul Rahman Mostafa Kotb El-Banna
Test \$ ✖	bt-test-arxlevant-12 ✖	Add note	105	Cairo	01129789595	ahmbela95@gmail.com	Ahmed Belal Abd-Allah Ali
Test \$ ✖	Assign	Add note		Alexandria	01112421335	amrosharkas@gmail.com	amr sharkas
Test \$ ✖	bt-test-arxlevant-09 ✖	Add note	111	Cairo	01156848304	asmaataha93.at@gmail.com	Asmaa Taha Shabaan
Test \$ ✖	bt-test-arxlevant-26 ✖	Add note	109	Cairo	1005065637	aya.egy11@yahoo.com	aya waheed
Test \$ ✖	bt-test-arxlevant-19 ✖	Add note	112	Alex	01099797699	hagarhosni62@gmail.com	Hagar Hassan Salem Hosni
Test \$ ✖	Assign	Add note	101	Alex	01226992992	Hekmat.hamdi95@gmail.com	Hekmat hamdi
Test \$ ✖	bt-test-arxlevant-16 ✖	Add note	102	Alexandria	01284557625	mazenbarakat1@gmail.com	Mazen Ahmed Barakat
Test \$ ✖	bt-test-arxlevant-01 ✖	Add note	103	Cairo	01146319161	tantawy1995@yahoo.com	mohamed magdy abdelrahman
Test \$ ✖	Assign	Available to work 5 to 6 hours, (4 hrs Fri&Sat & Saturday) 500 utterances	110	Cairo	0303374 0122	monahamed153@gmail.com	Mona Hamed Abd El

The applicant gets notified whether he has passed the test or not. When the applicant passes he gets the link and the password to his own account of the platform that they are going to use for Data Entry. Our system calculates the number of records entered, and convert this to payment according to the payment system. Whenever a new chunk of work is assigned, the applicant gets notified of his/her assignment for the next few days.

The screenshot shows the Liguistix Tank dashboard for user 'amr sharkas'. The main content area is titled 'مادة التدريب' (Training Material) and displays a list of transcription tasks. Each task includes a title, a duration, and a blue 'عرض' (View) button. The tasks are as follows:

المادة	عرض
1 - كيفية الدخول إلى حسابك	عرض
*1 - الدخول على البرنامج وبدء العمل	عرض
10 - القواعد العامة للكتابة	عرض
11 - الأمثلة	عرض
2 - بدء العمل	عرض
3 - الحالات العادية	عرض
4 - حالات التخطي	عرض
5 - حالة تحايل الأصوات	عرض
6 - حالة كتابة الموسيقى	عرض
7 - حالات عدم فهم الكلام رغم استنتاج حروفه	عرض
8 - التطينف	عرض

The right sidebar shows the user's profile 'amr sharkas' and a navigation menu with 'Transcription' selected, and 'حسابي' (My Account) and 'مادة التدريب' (Training Material) visible. Below the main content, there is a footer with the text '© 2019 Liguistix Tank'.

The screenshot shows the Liguistix Tank dashboard for user 'amr sharkas'. The main content area is titled 'الاجابة النموذجية' (Model Answer) and displays a list of audio files for transcription. Each file includes a title, a duration, and a play button. The files are as follows:

الملف	الاجابة النموذجية
0:03	5678
0:05	ثلاث أنواع الأول غني والثاني جميل والثالث فقير
0:02	تفجير ساحة تمانيّة
0:04	Skip
0:05	أبغى أقوى الدنيا كلها أقوى واحد في الدنيا كلها
0:08	أبي شباب اليومب أبي شباب اليومب أبي شباب اليومب
0:03	راح تصدى الساعة سبعة
0:03	Skip
0:03	أبغى مشعل و
0:03	أونور 5x
0:07	أجمل مجلس باللون تكوازي
0:03	أبغى أبغى

The right sidebar shows the user's profile 'amr sharkas' and a navigation menu with 'Transcription' selected, and 'حسابي' (My Account) and 'مادة التدريب' (Training Material) visible. Below the main content, there is a footer with the text '© 2019 Liguistix Tank'.

## Final Result

The final result gave the company exactly what it needed. a very simple solution for training and hiring suitable applicants. After implementation, the system was able to exclude around 10,000 from 12,000 applicants with no effort from the Linguistix tank team. As for accepted agents, they can go back to the training materials whenever they want to review them. The system also provided an adequate way for agents to check their progress, and get notified with new updates.

