Binaural Sound to Enhance Listening Comprehension in Language Tests

Teresa Magal-Royo (✉)
Universitat Politécnica de València, Valencia, Spain
tmagal@degi.upv.es

Jesus García Laborda
Universidad de Alcalá, Madrid, Spain

Abstract—The use of new digital formats in language learning and testing improves both the learning and acquisition skills development process of language tests. The process of listening comprehension is considered one of the most complex in the field of Computer Aided Learning Language, (CALL) because it relates to multimodal learning channels and brain sound perception in an unfamiliar communication environment (in a second language learning virtual environment) for learner. This article stresses the possibilities of using binaural sound in the design and implementation of tests as well as the cognitive issues that might be involved in the process of learning and assessment of a foreign language.

1 Introduction

Listening in a foreign language is a daily activity for most people around the world. It also an essential sense to perceive the world around us. Additionally, it is a critical matter in foreign languages (FL) where the communication environment is usually unfamiliar to the speaker and even more when the listener's FL competence is low.

Listening is part of a multimodal process where all senses participate. It can combine different perceptual modes like listening, viewing and body perceptual and corporal movement to capture the signification of the conversation. More importantly, if we focus on listening in a new language, the process is more complex than merely perceiving sound, because knowing a language also involves other skills such as grammatical use, pronunciation and the understanding and apprehension of the scene that which are instrumental supports in comprehension in a foreign language. Besides, interactions usually take a longer interactive time (since we subconsciously take time to translate or, at least to react. Thus, multisensory support facilitates communication.

When listening, interlocutors go through a complex process where they interpret the sounds to their acquired knowledge which is stored in their brain in order to be

used after first for comprehension and for production after. It is therefore the corner-stone of oral communication. Listening activities during the learning process are sometimes conscious and at other times subconscious, but in both cases the main information remains in our minds since our childhood or it is maybe in in our genetic information. Listening helps to understand and how the major foreign language environment functions but also rules, behavior and speech in different contexts. In Second Language Acquisition (SLA) it is very important to perceive and recognize listening situations happening in other countries and cultures for purposes that go beyond understanding, answering or arguing in another language. Communication helps to understand most sensations we get from real world environment, such as sounds that come from different sources, speakers' positions, one's own sound objects, musical instruments, etc.

This paper describes how binaural audio could enhance perception in a listening comprehension test in foreign language learning, by recreating the same audio conditions in the real world and how to implement it in the design of this type of tasks in an aptitude exam.

2   The Act of Listening

In second language acquisition, listening activities have emerged as an important component in the learning process [8]. Listening activity is a mental process which is hard to describe since the learners have recognize sounds, associate them to previous knowledge through understanding vocabulary and grammatical patterns, consider the context, comprehend the intention and finally re-interpret in their minds what on the situation and the message the interlocutor intends to transmit. From the academic perspective, this leads to an accurate interpretation and application or the capacity to apply what has been heard to academic tasks such as drills, exercises and, ultimately, tests.

For Rost, listening, as a whole is a:

"Process of receiving what the speaker actually says (receptive orientation); constructing and representing meaning (constructive orientation); negotiating meaning with the speaker and responding (collaborative orientation); and, creating meaning through involvement, imagination and empathy (transformative orientation)".

Widdowson [17] defines listening as the ability to understand how a particular linguistic chain is connected to everything is said in a particular act of communication. At this stage, the listener selects what considers relevant to their purpose and discards the irrelevant [17]

According to Acosta [1], the listener associates both linguistic knowledge such as well as non-linguistic such as the experience and knowledge of the world. Therefore, through inference the listener will not only be able to decode words, phrases and meaningful units already listened but also to interpret the sounds in their environment and context which helps to understand and interpret the entire communicational message.

This happens when language learners have to understand a conversation with a native and a non-native person of a language and have to use the lexical and grammar knowledge in a specific context real or created for assessment of a language [1].

3   Scene Sound Perception

One aspect to consider during the process of learning a foreign language is how we listen and learn from authentic situations and how we process the neuronal information in order to remember them and to use them later on. In general, scene perception is perceived through the multimodal integration of multiple sensations. Multisensory integration occurs in different parts of our brain. The process related to perceiving the audio scene is named Auditory Scene Analysis (ASA) and it was defined in the nineties as a group of cognitive processes combined in our mind, creating schemas. These schemas are used by the brain during the native and/or second language acquisition process to create and remember grammatical structures used in a social communicational context [3].

The latest research points out that diverse sensory systems are complementary among one another in a cross modal binding integration in the brain. Crossmodal binding integration is a term applicable to many phenomena in which one sensory modality influences task performance or perception in another sensory modality [2].

The best known is the process involving images and sounds signals combined in the real world environment recognition. This is especially important in language testing because when you perform a task for listening assessment, sometimes an audio visual source is used to improve not only the sound perception of the scene but to enhance and make real the act of communication which is intended to be produced in the very same input because the images also provide complementary information about the situation [16].However, in terms of cognitive constancy and remembrance the use of audio recordings in listening assessment provides better and more durable information than audiovisual recordings that combine audio and image.

There seem to be at least two types of brain mechanisms that take part when the brain integrates an audio scene with pertaining information. The first is a set of bottom-up processes which we probably share with other animals such as dolphins and others, [4]. Schemas he second type of mechanism used by our brain deals with audio frequencies and significant patterns in the environment which are considered innate. Large-brained beings such as primates, dolphins and other mammals like humans usually transform these while developing new ones through learning. Through this continuous process we are also able to get to recognize sounds, phonemes, words, patterns and sentences. Even songs and other sound chains which, in turn, permit the recognition of messages. These early or late processes related to Auditory Scene Analysis (ASA) are important in learning a new language through pattern recognition in conversations, situations that occur recurrently. There are two main theories which are related to the use of interactive learning: the innate and the statistic. It is our interest which mostly states that we acquire a language through the repetition of similar patterns a number of times. For instance, in order to acquire a sound through listening,

we have to hear one or a set of sounds. This way, we learn language patterns that help us understand the syntax, the grammar and the social communication of a language and generate new knowledge based on previous situations [12].

4   Binaural Audio Technology as a Multimodal Enhancement in the Listening Comprehension

The term binaural literally means to hear with two ears taking into account the sound perception in the brain based on the distance between ears. This happens in a dichotic listening when each ear receives different stimulus (usually in speech). The stereo sound for instance does not belong to dichotic listening because the stimulus sound is perceived at the same time of both ears and there is not a natural ear spacing. However, a dichotic perception stimulates the human perception to locate the origin of a sound by identifying cues from one ear (monaural cues), and by comparing cues received with both ears (difference cues or binaural cues). Among the cues, there are time slight differences of sound arrival and intensity differences (see Figure 1).
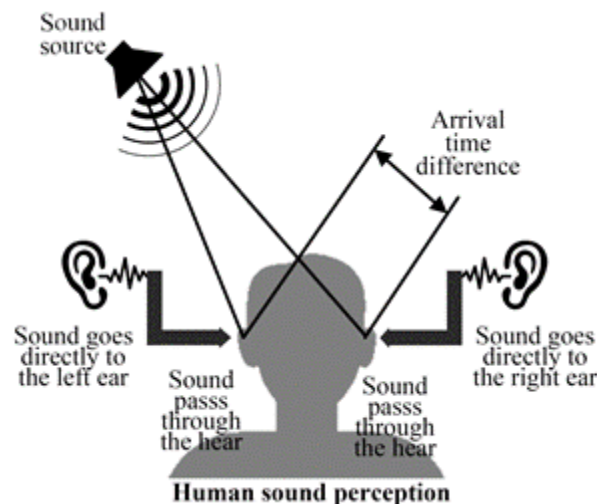

Fig. 1. Human sound perception

Wikipedia describes the process as it follows:

"The monaural cues come from the interaction between the sound source and the human anatomy, in which the original source sound is modified before it enters in the ear canal for processing by the auditory system. These modifications encode the source location, and may be captured via impulse response, which relates the source location and the ear location [5].

This impulse response is named the Head-Related Impulse Response, HRIR. Convolution of an arbitrary source sound with the HRIR converts the sound as if

it has been played at the source location, with the listener's ear at the receiver location. HRIR has been used to produce virtual surround sound [11].

The Head Related Transfer Function, HRTF, can also be described as the modification of a sound from a direction in free air to the sound as it arrives at the eardrum. These modifications include the shape of the listener's outer ear, the shape of the listener's head and body, the acoustic characteristics of the space in which the sound is played and so on. All these characteristics will influence how (or whether) a listener can accurately tell the direction from which a sound is coming [7]."

The method of recording binaural sound uses two microphones in a mock head with two ear-shaped pinnae, simulating the ear's natural position, creating an ear spacing or "head shadow" from the head to the ears. In fact, the student perceives Interaural Time Differences (ITDs) and Interaural Level Differences (ILDs). Adjustments (known as Head-Related Transfer Functions (HRTFs) are done automatically by the system itself. That means that sound is sequenced naturally as sound "involves" the human head and is adapted to the form of the outer and inner ear. The most often used commercial device is 3Dio Space Pro© developed 3Dio Company.

Fig. 2. 3Dio's Free Space© device.

5   Listening Comprehension in Second Language Acquisition (SLA)

As stated, listening comprehension is an interactive, interpretive process that links prior knowledge and linguistic interpretation from auditory oral and contextual segments of sound (usually as messages). Speakers use more their interpretation capacity or linguistic message according to each individual situation and especially the space and time contextual clues (consider irony intonation which is an issue in advanced language learning). Additionally, familiarity with the topic or conversational goals also have a role in listening comprehension as well as quality sound which differs significantly in real life and labs. For example, listening for overall general comprehension may not be as demanding (top-down processing) as listening for specific information where quality sound is a must (bottom-up processing) [15].

In general, there are two processes involved in listening comprehension:

- Listeners' dependence of background knowledge of the topic in transmission to comprehend the content of a message which includes the context, the conversation-type, cultural knowledge or other information stored in long-term memory as schemata (typical sequences or common situations) (top-down processes).
- Listeners' try to find out the main tokens (words) and contextual clues in the recording to create a tentative context of the communication. In this case, the listeners also use linguistic knowledge to understand the meaning of a message (bottom-up processes).

When listeners recognize the context of a text or an expression of the conversation, the process is way easier, because listeners can activate prior knowledge and make the appropriate inferences essential to understand the used audio message [6].

Therefore, it is necessary to prepare specific types of audio scenes to help students organize their judgements for the effective activation of appropriate background knowledge for general understanding, improving predictions and, most importantly, to prepare for listening. This significantly reduces the burden of comprehension of an audio scene for the listener when taking an exam or test.

Teachers should facilitate their students' capacity to understand their capacity to analyze and sequence their thoughts students, to stimulate appropriate related knowledge for understanding and making significate predictions, in order to prepare for listening. Through this facilitation process, students are capable to understand better but it also requires an excellent quality of input which, unfortunately, is missing in just too many occasions. On the other hand, listeners must reach an ability to listen selectively according to the purpose of the task, the type of listening required and the involved listening task strategies.

As O´Malley and Chamot (1989) point out, listeners use metacognitive, cognitive and socio-affective strategies to facilitate comprehension and to make their learning more effective. [13]. Thus, as far we know language learners require the following conditions to regulate their comprehension:

- Being able to recognize the referents

- Pay attention (just hearing is not enough to understand)
- Focus on comprehending the message predictions
- Self-monitor the input
- Use metacognitive knowledge

But they also require:

- Good comprehensible input (quality of listening)
- Capacity to distinguish sound chains
- Avoid communication noise (which is not just outside noise but is also induced by the echo of stereo sound through traditional headphones)

Test administrators and teachers should aim at getting the best possible sound. This paper enhances the importance of binaural sound because it meets these three aspects and thus make listening more accessible and also favors the student's use of personal strategies of learning. This is especially important because in testing interferences should not compromise the test taker's performance. This is especially valid for online testing since communication or the net sometimes can have negative side effect that must be avoided.

Therefore, dual sound in language tests requires the use of educational digital platform to prepare and assess for which students must become acquainted.

For practical purposes, the listening test assessment of a foreign language has restrictive conditions given the following aspects:

- Technological availability, knowledge of the environment and / or the virtual platform, test duration, conditions of the transmission medium.
- Using a multimodal interface to enhance the conditions of input and output data during the task involvement.
- Formal tasks according to the same test standard for evaluating listening comprehension skills that must adapt to technology criteria available including different types of communication.
- Development of a digital stage interface or an interface appropriate to the needs of the tasks, user-oriented and accessible by offering an easy to understand environment. This aspect, although it could be included within the technological section, should be treated as a crucial aspect including its conceptual and visual development focused on helping users stay comfortable during the test [10].
- Digital content creation of listening situations according to social and cultural environments of the users and the previous knowledge will be evaluated. In that case, simulating the normal users audio conditions similar to the environment in which they learned or developed their listening comprehension

6 Binaural Audio Implementation in a New Design Scenario for Listening Comprehension Assessment Test

By implementing binaural audios we can improve the sound experience and the perception of a real auditory scene that helps test takers (and also foreign language learner)to understand a common scene or situation in close-to-real real world and therefore improve the multimodal learning conditions involved in foreign language learning The implementation of a multimodal learning process using binaural audios of conversational or acting scenes facilitates and enhances the cognitive listening learning conditions in a foreign language.

In the case of a listening comprehension, test operation through the use of binaural audio allows the user to focus more on a perceptive process for direct interpretation of the reconstructed auditory scene, and additionally, use language testing strategies which are also developed while learning their mother tongue but reframed and increased in the process of learning a foreign language [9].

The design of the test environment must take into account the use of a multimodal interfaces during the navigation process that can be previously selected by the user. To implement a binaural audio into assessment tasks in a listening comprehension test with a multimodal interface we must also control the process of data input and output that the user has to control in order to perform the specific task

Using a tablet Pc may also offer to the student a selection of, at least, two types of navigation supported in a conventional web browser (touch, voice recognition and mouse) for specific navigation processes during the task (touch, voice recognition, mouse and keyboard entry) and to solve and fill out the test questions.

There must also be a verification procedure with a binaural sound test to detect auditory or physical problems with the cue synchronization. Sometimes the student takers will not perceive the audio and therefore, they cannot perform the test properly.

Finally, in relation to the listening task, test designers should develop items considering the specific test rubrics, the test delivery facilities, the type of on-line test questions (such as those that can involve more than one transmission mean and the question´s text on recognition of the spatial location of sounds, objects or people involved in the binaural audio scene.
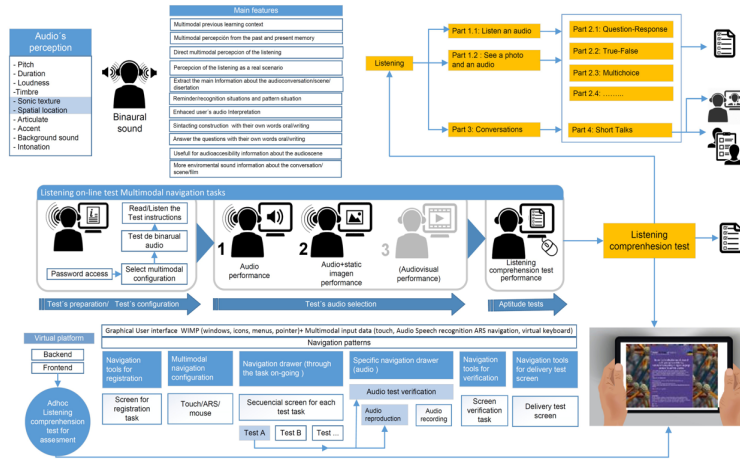
Fig. 3. Technical approach to Listening comprehension test scenario with binaural audio

7   Conclusion

The use of binaural sound for a listening comprehension assessment is achievable and useful since it enhances the multimodal learning experience during foreign language learning and testing. The implementation of binaural sound for an auditory scene user´s recognition can help both user´s recognition as well as the teacher to have better tools during the process of language learning as well as in engineering introductory projects [18, 19] for cognitive improvement [20]. As a consequence, binaural sound Binaural audio provides a cognitive perception of the global scene in an immersive way that will extend the creation possibilities of specific listening testing. The design and implementation conditions of binaural audios in a listening tests, as it has been stated, improve perception sequencing that will lead to a better performance in oral aptitude tasks. In listening tests design not only aspects related to the implementation of the sound must be considered but also the use of a multimodal interfaces that directly help the user during the test administration, delivery and performance.

Multimodal interface creation for that purpose allows user´s control of the input and output adapted to some specific tasks of the listening comprehension process.

In conclusion, the use of these types of audios will allow to enhance the performance in tests and decrease the cognitive load during the test performing process, reduce anxiety due to recording noise. This partly because students receive the multimodal communication directly into the brain and the cognitive process will be quickly understood as a real listening.

To conclude, binaural sound has an enormous potential in language tests that has been already discussed. However, this short paper just outlines the prospective issues in its use in task and test design. Further research is required since few papers have reported real student data so far. Future work will have to focus on performance, stu-

dents' attitudes, facilitation and cost of design and delivery. In this sense, this paper has only intended to show a theoretical approach that soon will be followed by further experiments.

8 References

[1] Acosta, R (1984). Communicative Language Teaching. Sumptuous Publications, Altman.

[2] Bizley J K., Jones G P & Town S M, (2016). Where are multisensory signals combined for perceptual decision-making?.Current Opinion in Neurobiology, Vol 40 pp.31–37 https://doi.org/10.1016/j.conb.2016.06.003

[3] Bregman A.S. (1994). Auditory scene analysis: The perceptual organization of sound. Bradford Book, London, ISBN: 9780262521956

[4] Bregman, A. S. & McAdams, S. (1994). Auditory Scene Analysis: The Perceptual Organization of Sound. The Journal of the Acoustical Society of America.Vol. 95 (2), pp.1177-1178. https://doi.org/10.1121/1.408434

[5] Brungart DS, Rabinowitz WM. (1999). Auditory localization of nearby sources. Head-related transfer functions. Journal Acoustic Society of America Vol. 106 (3), pp.1465-1479. https://doi.org/10.1121/1.427943

[6] Byrnes, H. (1984). The Role of Listening Comprehension: A Theoretical Base. Foreign Language Annals Vol. 17 (4) pp. 317-29. Cambridge: Cambridge UP Series: Language in use. https://doi.org/10.1111/j.1944-9720.1984.tb03235.x

[7] Dunai L.,Peris-Fajarnés G., Magal-Royo, T., Defez B. & Santiago Praderas V. (2011). Virtual Moving Sound Source Localization through Headphones in Advances in Sound Localization, book edited by Pawel Strumillo, InTech Publisher, Croatia ISBN 978-953-307-224-1. https://doi.org/10.5772/14643

[8] Feyten, C. M. (1991). The Power of Listening Ability: An Overlooked Dimension in Language Acquisition. The Modern Language Journal Vol.75 pp. 173-80. https://doi.org/10.1111/j.1540-4781.1991.tb05348.x

[9] Magal-Royo T, Garcia Laborda, J. and Price S. (2017). A New m-Learning Scenario for a Listening Comprehension Assessment Test in Second Language Acquisition (SLA). Journal of Universal Computer Science, Vol. 23 (12) pp. 1200-1214.https://doi.org/10.3217/jucs-023-12-1200

[10] Magal-Royo T. & García Laborda, J, (2017) Multimodal Interactivity in Foreign Language Testing. In Multimodal Interaction with W3C Standard. Toward Natural User Interfaces to Everything, Chapter 19. Dahl, Deborah (Ed.). Springer International Publishing. ISBN 978-3-319-42816-1. https://doi.org/10.1007/978-3-319-42816-1_16

[11] Møller, H., Sørensen, M.F., Hammershøi, Dorte; Jensen, C.B. (1995). Head-Related Transfer-Functions of human subjects. Journal of the Audio Engineering Society, JAES Vol.43 (5) pp. 300-321. http://www.aes.org/e-lib/browse.cfm?elib=7949 Retreived 12/05/2018

[12] Nakaizumi, C. Mori, K.Matsui, T., Makino S. & Rutkowski, TM. (2013). Auditory Brain-Computer Interface Paradigm with Head Related Impulse Response-based Spatial Cues. 9th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS) pp. 806-81 https://doi.org/10.1109/sitis.2013.131

[13] O'Malley, J. M., Chamot, A. U. & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. Applied Linguistics Vol 10(4) pp. 418-437 https://doi.org/10.1093/applin/10.4.418

[14] Rost, M. (2013). Teaching and Researching: Listening. Routledge, London, UK. ISBN: 9781315833705

[15] Vandergrift, L. (1999). Facilitating Second Language Listening Comprehension: Acquiring Successful Strategies. ELT Journal Vol. 53 pp.168-76. https://doi.org/10.1093/elt/53.3.168

[16] Villegas, J (2015). Locating virtual sound sources at arbitrary distances in real-time binaural reproduction. Virtual Reality. Vol. 19 (3-4) pp.201-212 https://doi.org/10.1007/s10055-015-0278-0

[17] Widdowson, H. G, Cook G. & Seidlhofer B., (1995). Principle & practice in applied linguistics, Oxford University Press.ISBN: 978-0194421485

[18] Gero, A. (2012) Improving Intrinsic Motivation among Sophomore Electrical Engineering Students by an Introductory Project. International Journal of Engineering Pedagogy. Vol. 4(2) pp.13-17. https://doi.org/10.3991/ijep.v2i4.2247

[19] Gonzalez Hernandez, W. (2013). Creativity Development in Informatics Teaching Using the Project Focus. International Journal of Engineering Pedagogy. Vol. 6(2) pp.4-13.

[20] Drigas, A. & Karyotaki, M. (2014). Learning Tools and Application for Cognitive Improvement. International Journal of Engineering Pedagogy. Vol. 4(3) pp.71-77.

9   Authors

Teresa Magal-Royo, PhD in Fine Arts, Associate Professor in the Graphic Engineering Department. Degree: Product Design. Universitat Politécnica de Valencia, Spain. Postgraduate in the Royal College of Art, London, England. Project coordinator related to the use of new technology in the field of education/Design and development of user-oriented interfaces/ Adaptation of graphic interfaces to mobile devices/ user experience in education/Product design and digital communication for business. /Multimodal interfaces user oriented for ubiquitous devices. She has participated in national and international research projects related to technology and education. Mail: tmagal@degi.upv.es http://www.upv.es/ficha-personal/tmagal
Researcher ID: K-6728-2014 ORCID: https://orcid.org/0000-0002-7640-6264

Jesus García-Laborda PhD in English Philology, Universidad Complutense de Madrid, UCM, 2000, DC Doctor of Education (Teaching of Language and Literature). European mention. UCM. 2010. Master's in ESL - U. of Georgia, 1992. Master's in Language and Comparative Literature - U. Wisconsin. Currently works as Dean of the College of Education of the University of Alcala de Henares. He teaches Didactics of English as a Foreign Language. He is Academic bureau member of Franklin Institute. His research is related to Applied Linguistics in the field of Computer-assisted Second Language Learning. He has participated in various national and international research projects on the use of technology in education related to learning of second language. Mail: jesus.garcialaborda@uah.es Researcher ID: E-1797-2012
ORCID: http://orcid.org/0000-00 03-0125-4611