

# Listening comprehension and sound image: The listening gap 「聴解と聴覚像—リスニングのギャップ」

Maria Gabriela Schmidt

シュミット、マリア・ガブリエラ, 人文社会系・CEGLOC

## 1 Introduction

Listening is more than a skill and it is as a teaching subject often undervalued in foreign language teaching (Field, 2008: p.1). Most teachers will rank it behind grammar, reading, speaking and writing. Language is in its very center sound and meaning based. The sound we perceive and the sounds we utter are related to meaning, even when they are not lexical words, such as the interjection “oh”, expressing some kind of surprise. This article starts with the assumption that the relation of sound and meaning is vital for building up listening abilities in a foreign language. First, we will look at research results of psycholinguistics on how listeners are primed by their own language acquisition. The mother tongue has to be taken seriously into account. In order to accommodate the research results with the teaching practices, F. de Saussure’s term “sound image” will be used here, to explain an overlooked gap in the foreign language teaching practice. The author suggests here, to name it a listening gap. Becoming aware of it over time during instruction, the listening gap is to some extent related to listeners “being unable to form a mental presentation from words heard” (Vandergrift & Goh, 2012: p.22) and to “make matter worse, the plight of the weak listener often goes unrecognized” (Field, 2008: p.3).

The listening gap in foreign language instruction opens up, because the listening exercises in classroom instruction follow the tradition of mainly written and grammar based input. But results from neuroscience show that “orthographic learning enters the brain in a different way from phonological processing”. (Tokuhamma-Espinosa, 2011: p.130) Without primary meaning-building sound input, the relation between sound and meaning is discontinued. The listening gap is used here as a term, to describe the phenomenon that learners do create their own “made-up sound images” for mainly written linguistic input through reading. If there is no sound available to express our thoughts, we may need other signs instead such as letters, pictographs, our hands or even our body. They may function as a replacement for sound. This paper will reason on the sound and meaning relation, describing the listening gap through sound images, and will argue for a more diversified extensive sound input, to develop listening as a real ability and give learners a chance to develop the sound-meaning relation for processing language successfully.

## 2 The role of the mother tongue

In teaching listening, the first thing to keep in mind is the mother tongue of the learners. This may be a very well-known general knowledge. But comparative studies involving researchers from various countries were conducted by the psycholinguist Anne Cutler. The findings are based on studies on language acquisition across a number of languages. She emphasizes (Cutler & al., 2007: p.106):

Listeners process spoken language in ways which are subtly tailored to suit the phonological structure of their native language. This has been best documented with respect to the segmentation of continuous speech into its component words. This draws on rhythmic structure, and because rhythm differs across languages, speech segmentation procedures are likewise language-specific. [...] The discovery that segmentation procedures are language-specific provided a potential explanation for the frequent experience that segmenting speech in a nonnative language (henceforth: non-L1) is hard. Where the rhythmic structure of the non-L1 differs from that of the native language (L1), use of the L1 processing strategy will be counter-productive. Experiments confirmed that non-L1 listeners did not necessarily segment speech in the same way as L1 listeners. ...Instead, listeners tended to apply their L1 listening strategies inappropriately to the non-L1 input. French use of syllabic procedures with English [...] and Japanese [...] was paralleled by Japanese application of moraic segmentation to French, Spanish, and English [...]. Following those studies, there have been many demonstrations of inappropriate application of L1 segmentation to non-L1 input, even when the non-L1 is a high-proficiency L2.

For a listener the continuous speech appears to be one piece of information, and it is hard to parse "the sound stream into meaningful units when word boundaries are difficult to determine because of stress pattern, elisions, and reduced forms". (Vandergrift & Goh, 2012: p.21) Here we find the relation of sound and meaning. Our ability to determine meaningful segments in perceived sounds is related to our mother tongue. The stress pattern of the target language is difficult to catch, because the stress pattern of our mother tongue is playing a crucial role and impedes perception. Taking these findings into account for teaching a foreign language, listening exercises should help to develop a sound bridge from one language structure to the other. To cope with the gap of perception, using top down metacognitive strategies for anticipating the listening contents may support the process of comprehension. On the pedagogical side, exercises should reflect the sound structure of the mother tongue and of the target language in order to build bottom up units of sound and meaning for the target language.

### 3 The terms “sound image” and “made up sound image”

As mentioned in the introduction, the term “sound image” is used and introduced here, to name a special phenomenon in foreign language listening, when the foreign language is taught in a classroom. In that case, a listening gap can be determined. This is related to foreign language instruction following the tradition of mainly written and grammar based input. The term “sound image” functions here as a tool to describe the process of listening to sounds and establishing meaning. In phonetics, this is called “representation” or “manifestation” (Niebuhr, 2007: p.2-3), referring to different states of a transferred sound pattern.

The author prefers to use the more traditional and linguistic term “sound image” as introduced by linguist Ferdinand de Saussure, who uses in his *Cours de linguistique générale* (Engl. translation *Course in general linguistics* by R. Harris 1983), as reported by his students Ch. Bally and A. Sechaye, not “sound” which would be related to the physical sound (wave), but “sound image” (in French *image acoustique*, in German *Lautbild*, in Japanese 聴覚像 *cyou kaku zou* or 音の表象 *oto no hyouzou*) as an imagination of sound. A linguistic sign (word, expression) consists therefore of its meaning (*signifié*) and its related sound image (*signifiant*), as two sides of a leaf. This double relation is the basic assumption for this approach. The sounds we perceive when listening are not the sounds the other person uttered. Also, conversely, the sounds we have in mind when speaking are not the sounds we produce, and are not the sounds the other person perceives. This may sound trivial, but it is not. It has to be taken seriously and be considered in teaching listening. Niebuhr (2007) emphasizes that sound transferring the information from the speaker to the listener as given in the basic model of communication changes several times in its carrying medium and adapts every time in a special manifestation to this medium (Niebuhr, 2007: p.2-3):

[concept of utterance]  $\leftrightarrow$  (1) speaker prepares for speaking (nerves)  $\leftrightarrow$  (2) muscles of speech organs of the speaker move  $\leftrightarrow$  (3) sound waves are transmitted through air  $\leftrightarrow$  (4) listener receives sound through the ears (body-mechanical)  $\leftrightarrow$  (5) transformed into signals (nerves)  $\leftrightarrow$  [perception of utterance]

The sound image is a generated representation of a sound or a sound pattern. This fits very well with the findings of neuroscience that meaning and sound are established through pathways between neurons in the brain. The one side is the intended meaning by the person uttering something and the other the meaning, the person listening, is constructing by the perceived sound. “A speaker does not implant a message in the learner’s mind. The listener has to remake the message” notes Field, and stresses that the “message is a product of the individual listener” (Field, 2008, p.6). This can be visualized by the following model:

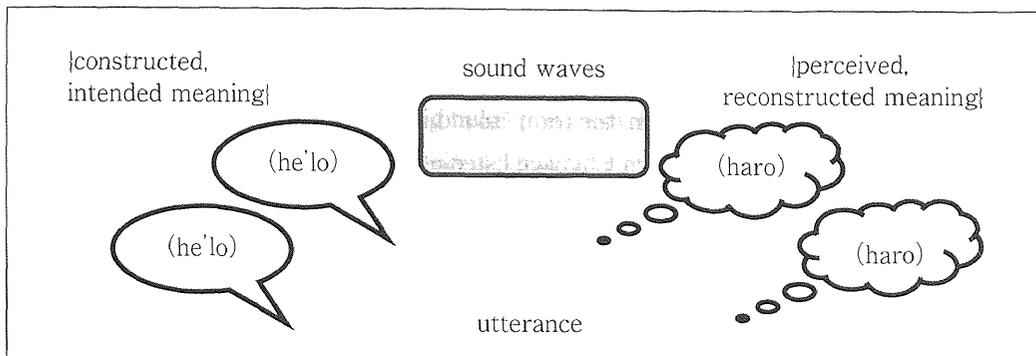


Figure 1: Manifestations of sounds

During classroom instruction when some audio input is given, for example the greeting “hello”, we would find as many representations as students are in the class, plus the one of the teacher. To make sure that there is no ambiguity at all, the teacher will write down the sentence, and repeat it several times. In teaching listening, the relation between meaning and sound image is crucial, but it cannot be controlled that easily. There is no way to make sure that the sound image of the learner and the intended sound image do match. A mismatch is therefore very likely. To give an example, Table 1 shows a very simplified scheme:

teacher					student
thinks	“he’lo”			“?he’ro, ?hiro?”	thinks
says	/h3’lo/			/hi’ro/	hears

Table 1: Production and perception of sounds

The teacher wants to introduce an expression for the greeting “hello” and says /h3’lo/. But the student perceives /hi’ro/ instead and is not sure about the meaning. The students may interpret it as ‘?he’ro’ or ‘?hiro’ and may not come to any conclusion, because the meaning does not fit the situation. This example may be very simple, but the surprising reaction of students during classroom instruction to seemingly simple expressions has to be taken into consideration. Vandergrift & Goh (2012: p.21) state that

word segmentation skills are language-specific and acquired early in life. They are so solidly engrained in the listener’s processing system that these segmentation strategies are involuntarily applied when listening to a non-native language.

The top down procedure to find a match with a strategic guess to cope with the continuous incoming sound stream, here to anticipate a greeting, and the bottom up procedure, when the incoming sound patterns do not match already established lexical items, may result just in giving up. Field stresses that the needs of weak listeners are often not recognized. This happens

“partly because of the inaccessible nature of listening, which can only be tested indirectly, often by means of cumbersome comprehension questions” (Field, 2008: p.3). He continues that learners

sit in class day after day, comprehending little and often blaming themselves. But the fault lies with a system that does not accord sufficient attention to the skill that, above all others, is crucial to their learning. (Field, 2008: p.4)

The learner may construct a sound image on the basis he or she perceived the sound, and how it makes sense to him or her. The author suggests naming these sound images ‘made up sound image’. Especially when the teaching content is grammar and reading oriented, the learner will more likely have a visual image instead of a sound image. The reason is that he or she has not had enough sound input to build up sound images for various linguistic units. Sometimes these made up constructions are used for a long time, hardly detectable and without any chance of mending, “even when the non-L1 is a high-proficiency L2” (quotation from Cutler above). When the made up sound images for the target language are established, the learner may experience difficulties when listening to authentic speech. Influenced by their own native language AND the made up sound images he or she has in mind, the sound perceived of the target language will not fit well the circumstances of the utterance. A likely consequence would be frustration and so-called listening anxiety. Field stresses that listening is not tangible in the way that speaking and writing are” (2008: p.1), and it is not easy to achieve demonstrable results, because of the lack of measurable benefits. Teachers think,

Once the learner’s ears have adjusted to the phonology of the target language (the argument goes), listening skills from the first language (L1) will transfer themselves to the second (L2) by some process of osmosis. (Field, 2008: p.2)

Listening as a teaching subject is neglected, because teachers and learners think, that it is acquired naturally, by the way of doing all other language activities (reading aloud, listening to the teacher’s voice, doing conversation exercises, listening to the CD). But in order to address this phenomenon properly, foreign language teaching has to take a totally different approach. Teaching listening should not be exercises for testing listening, not vocalized written texts or spoken conversation, but should start to use sounds and sound input to build up meaning, meaningful linguistic units, linguistic units which can build a realistic base for the bottom up comprehension strategy.

## 4 Teaching Listening

Among the publications on teaching listening in recent years, two studies are widely acknowledged as shedding more light on this topic. One is *Listening in the Classroom* (J. Field 2008) and the other *Teaching and Learning Second Language Listening* (L. Vandergrift and C. Goh 2012). Both publications advocate listening as an active meaning building process and promote a more learner centered approach in the classroom. The latter discusses metacognitive strategies to develop better listening skills. This seems to be a very promising step forward considering the classroom pedagogy for listening.

### 4.1 The linguistic gap: Two aspects of massive sound input

First, in our mother tongue, we are in a continuous stream of sound input, not only linguistically. Sounds are part of our daily interaction. When you open the door, you remember a specific sound. This will evoke memories and emotions instantly. When using the word 'door', the linguistic meaning may be 'something that opens up and you can enter a room or a building'. But individually, everybody will hear and see a different door, depending on his/her experience. The storage of meaning in our brain is related to the sound image. Therefore sound is related to meaning and emotion in relation to our experience. More than reading, sounds evoke immediate emotions. We even may judge the person speaking as likable or unlikeable after the first words uttered. (K. Reinke, 2015; Shaules, 2015; I. Berent, 2013) As for a foreign language learner, the sound input is mainly based on the sounds provided during classroom instruction, the voice of the teacher and for some seconds, a recording related to the teaching material.

Second. When we talk about teaching listening, we have to not talk about a physical chain of sounds, but of images of sound as representations, which are different for everybody taking part in a communication situation, as mentioned before. In our mother tongue, we are normally trained to cope with disturbances as loud background noise, elisions and errors. Based on our massive linguistic sound input and our language experience, we can mend the missing parts from various angles and understand the message. (Rickmeyer, Sichelschmidt & Strohner 2005) But as a foreign language learner, we do not have enough linguistic input or experience to use the language in that way, yet. So we have to rely on the acquired knowledge like vocabulary or grammar or the strategies to cope in classroom and in tests with hard to understand passages.

These are the two sides of the linguistic gap. The sound input is not continuous and cannot create vital linguistic units to build a base of meaningful sound units, which can be processed automatically. Linguistic units are mostly built from reading, writing, learning grammar rules, and with made up sound images.

## 4.2 Other approaches

Considering the influence of the mother tongue, Shimaoka (2014) uses an approach to adapt the pronunciation of the target language. He suggests that American English speakers should pronounce his first name “Takashi” as “Taxi”, because this is probably close to the sound pattern they may perceive. On the other hand he advises Japanese students of English to look for sound similarities even if they are unusual: “Mind the door” and read it 「満員だぞー」 (*man in da zou*), meaning “All booked out.” Or, “Come here” 「亀部屋」 (*kameheya*) meaning “turtle room”. This approach may be unusual but our brain looks for patterns and similarities and learns easily, when they are new and surprising. (Tokuhama-Espinosa, 2011: p.212) Looking for an unconventional pedagogical approach may help to overcome the aforementioned listening anxiety and help the students to get some confidence.

For controlled and automatic processing during listening, it is important not to overemphasize rules and vocabulary, which could overload the working memory.

“While long-term memory shapes the interpretation of what listeners hear, working memory influences the efficiency of the cognitive processing and allows the listener to think about appropriate response, as in the case of interactive listening.” (Vandergrift & Goh, 2012: p.20)

Integrating a more affective side of the listening exercises with emotionally relevant examples can prepare the students' ears for more cognitive, challenging tasks. (Tokuhama-Espinosa, 2011: p.213) Using films (audio-visual material) for teaching listening is a very promising way (Fujita 2015). If students do not feel up to the speed of the authentic material, it is important to prepare by using strategies as anticipating hypothesis, predicting the story, to think of possible expressions, to watch the film in short segments and to verify in small groups the hypothesis. The approach of Vandergrift and Goh (2012), of focusing on comprehension strategies can be a very useful tool to cope with listening tasks. Another, more phonetic-linguistically informed approach for opening the listening ears from the mother tongue to the target language, is a step by step approach taking into account the linguistic structure of both languages. (Schmidt 2015)

The author wants to express the following assumptions related to listening:

(1) Learners need sound input to create sound images, in order to process language. When there is no adequate sound input, learners will create in their minds so called 'made-up sound images'. The constructed sound images may obstruct the perception of the real sound.

(2) The term “sound image” by Saussure seems to have more depth and relevance, including the established relation between sound and meaning. It should be considered and used in the discussion of teaching listening, as the term representation is used in a more phonetic sense.

(3) Listening input should be diversified and below the actual level of the learner. A large amount of audio input is needed to develop appropriate sound images.

(4) Pronunciation should be supported by body movement, including prosodic patterns for communication units, chunks, sentences, and giving every utterance a proper situational sound-meaning relation. Pronunciation exercises can support listening and vice versa.

More detailed studies are needed on the aspects mentioned. Especially on the side of sound perception, and how learners of foreign languages can be supported to work on their perception of sounds, and to adjust it closer to the target language.

### 5 Teaching methods for listening in the classroom

Teaching listening should take seriously into account the above mentioned aspects. There may be others, but these are the basic ones. The perception of sounds in the classroom is different from the perception of sounds in a natural social environment. Here are six suggestions for fostering listening in the classroom which are partly inspired by neuroscience (Tokuhamas-Espinosa 2011; Schmidt 2015):

(1) Developing	Mother tongue	Target language
Developing step by step listening exercises to become accustomed to consonant clusters	Japanese: Moraic structure resulting in Katakana- English st- : su-ta; str- : su-ta-ra	English: Consonant cluster: <u>s</u> and, <u>t</u> ent / <u>s</u> tand / <u>s</u> trand

Table 2: Developing phonetic structure step by step

The phonetic structure including intonation, word stress, prosody has to be taken very seriously. The teacher should anticipate and analyze possible difficulties in detail and try to develop strategies and exercises to help learners to cope.

(2) Relation of	Meaning	Sound
Use sounds as 'audio flash cards' - Create various 'audio flash tracks' (different voices)	Establish meaning by connecting sound pattern	Listen to a sound pattern
Play 'audio flash tracks' - let students look for cards or things related to the sound	Establish meaning by connecting sound pattern	Connecting a sound pattern

Table 3: Relation of audio flash cards

Flashcards are a very helpful tool to overcome/remember written input. Therefore in a pedagogical adaption “audio flash tracks” are needed as well, to create direct audio input and build up meaning without written or visual support.

(3) Bring	Sounds	In the classroom
	Let students bring sounds in the classroom and let them talk about it.	

Table 4: Bring sounds into the classroom

Get the students involved and let them bring sounds to the classroom. Most students have a cell or smart phone. Sounds are very easily used for teaching foreign languages using these tools. As sounds are related to emotions, this is a mighty tool using various findings of neuroscience (novelty, emotions, personalization). (Tokuhamma-Espinosa, 2011: p.213)

(4) Catch	the words and	the meaning
Introduce key words in various listening contexts and let students imitate.	Use texts and conversations of a former lesson, slightly below the level, possible for learners to catch some words.	Catch the meaning. Students can tell what had happened (in the mother tongue or target language, their choice!)

Table 5: Catch keywords below level

A written word can be read several times, a spoken word passes by. The target text for the listening exercise or the chunks for the conversation should be slightly below the actual level of the students. If they do not catch the word (sound), they might give up and be frustrated. Catching the meaning is related to the working memory. Learners gain confidence when they are able to catch the meaning. Let the student use the mother tongue to tell the understood parts of the story.

(5) Listen for fun	Stories	Conversations
Play listening input that is not related to any task. Just for listening – extensive listening, continuous listening.	Simple stories, stories the students already know	Simple conversations, with everyday situations that are easy to catch

Table 6: Listen for fun

The amount of audio input provided by textbooks and listening related exercises in classrooms is not enough to establish and develop proper sound images of the target language. Therefore extensive listening in the sense of a massive sound input is needed, not only intensive listening

for discrimination of sounds and words. Extensive listening (in the multiple sense of its meaning) is important for getting a sense for the sounds of the target language.

(6) Look up	Sound pattern	Prosodic elements
The perception and the production of sounds may interact continuously.	Sounds are mostly produced in meaningful combinations as words, expressions, utterances.	Prosodic elements do give clues how to evaluate the message inside the linguistic units

Table 7: Look up older textbooks for pronunciation

Concerning pronunciation, include the prosodic patterns of utterances and sentences. Look up older textbooks. They may well be better than we think they are, for example, H. Gordon Grate *English Pronunciation Exercises for Japanese Students* (1974).

The author suggests a more diversified, creative approach for listening in the classroom, to give students the opportunity to develop sound ‘sound images’. This is vital for all four language skills. However, concerning sound and meaning in teaching listening, the perception of sounds in classroom is different from the perception of sounds in a natural social environment. This is the reason why research results have to be discussed and cannot be transferred directly into classroom pedagogy.

This paper emphasized three aspects of listening. The mother tongue of the learner plays a crucial role for perception and production of sounds. For processing language we need sound images. If there is not enough appropriate sound input learners may use made-up sound images, which has not been addressed so far in academic studies. And finally, teaching listening has to consider the sound-meaning relation. To do so, a whole new set of listening exercises is needed to close the listening gap in foreign language teaching. This paper is a summary of a preliminary study to identify some aspects of teaching listening.

## Acknowledgment

*Results presented here are related to the research project on ‘Extensive Listening and Foreign Language Teaching’ supported by a grant of the Faculty of Humanities and Social Sciences at University of Tsukuba during the academic year 2014-2015.*

## References

- Berent, I. (2013). *The Phonological mind*. Cambridge University Press.
- Cutler, A.; Kim, J. & Otake, T. (2007). “On the Limits of L1 Influence on Non-L1 Listening: Evidence from Japanese Perception of Korean.” In: Warren, P.; Watson, C. I. (eds.). *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, p.106-111.

- Field, J. (2008). *Listening in the Language Classroom*. Cambridge: Cambridge University Press.
- Fujita, Ryoko (2015). The Effect of Using Films as Teaching Materials on Japanese EFL Learners' Listening Comprehension Abilities. (Tsukuba University, Faculty of Humanities and Social Sciences. Unpublished Ph.D.)
- Gordon Grate, H. (1974). *English Pronunciation Exercises for Japanese Students*. New Jersey: Prentice Hall Regents.
- Niebuhr, O. (2007). *Perzeption und kognitive Verarbeitung der Sprechmelodie. Theoretische Grundlagen und empirische Untersuchungen*. Berlin, New York: Walter de Gruyter. (Language, Context, and Cognition Vol.7).
- Reinke, K. (2015). Gesprächs- und Aussprachekompetenz im DaF-Unterricht – Grundlagen, Probleme, Perspektiven. In *Mündliche Kommunikation im DaF-Unterricht: Phonetik, Gespräch und Rhetorik*. Niikura, M.; Hayashi, R. et al. (eds.). München: Iudicium, p.37-64.
- Rickmeyer, G.; Sichelschmidt, L. & Strohner, H. (2002). *Psycholinguistik*. Tübingen: Stauffenberg.
- Saussure, F. de (1916). *Cours de Linguistique Générale*. Ch. Bally, A. Sechehaye (eds.). Geneva.
- Saussure, F. de (1983). *Course in General Linguistics*. Translated by R. Harris. LaSalle, Illinois: Open Court.
- Schmidt, M. G. (2015). Phonetik rezeptiv und produktiv im Bereich DaF in Japan – mit einer methodischen Überlegung. In *Mündliche Kommunikation im DaF-Unterricht: Phonetik, Gespräch und Rhetorik*. Niikura, M.; Hayashi, R. et al. (eds.). München: Iudicium, p.128-148.
- Shaules, J. (2015). *The intercultural mind. Connecting Culture, Cognition and Global Living*. Boston, London: Intercultural Press (Nicholas Brealey).
- Shimaoka, Takashi (2014). カタカナ活用 英語発音力向上の試み, 筑波大学名誉教授.
- Tokuhama-Espinosa, T. (2011). *Mind, Brain, and Education Science. A comprehensive Guide to the new brain-based teaching*. New York: W. W. Norton & Company.
- Vandergrift, L., & Goh, C. (2012). *Teaching and Learning Second Language Listening. Metacognition in Action*. New York: Routledge.