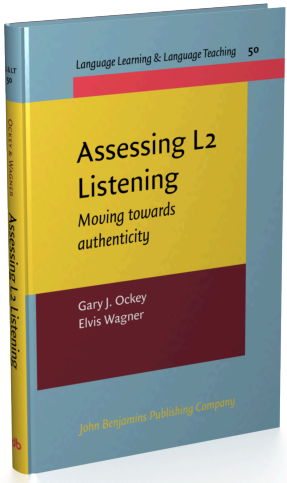
Language Learning & Language Teaching 50

Preface

https://doi.org/10.1075/lllt.50.preface



Pages xi–xvi of
Assessing L2 Listening: Moving towards authenticity
Gary J. Ockey and Elvis Wagner
[**Language Learning & Language Teaching**, 50]
2018. xvii, 278 pp.

© John Benjamins Publishing Company

This electronic file may not be altered in any way. For any reuse of this material written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

For further information, please contact rights@benjamins.nl or consult our website at benjamins.com/rights

Preface

In the 1980s, when I first became interested in assessing second language (L2) listening comprehension, listening was often referred to as the "forgotten skill." Despite some excellent work in the intervening years, in my experience, it is still the most neglected of the traditional four skills. This is unfortunate, because in a number of ways, listening can be regarded as the most fundamental skill of all. It is certainly the first of the skills developed by children when learning their L1, and there is considerable evidence, both empirical and theoretical, that L2 learning is greatly facilitated by an early emphasis on listening.

However, the reason I consider listening so important is because it helps develop automatic processing skills. Or to put it another way, it helps, or even forces, the student to internalize the core linguistic and cognitive processing skills that are so necessary for successful language use. Language processing is a very complex skill, and the human brain has two main ways to perform complex skills, whether physical skills such as driving a car, or cognitive skills, such as comprehending language. The first of these is through conscious controlled processing – thinking about how to do it, while doing it. The second way is through automatic processes that take place outside conscious control. Automatic processing is much faster and far more efficient than conscious, controlled processing. Of course, most complex skills start out as conscious processes, but then with practice, we develop the ability to do them with less effort, until eventually they become fully automatic, and outside conscious control. Language processing is far too complex to be done consciously, and successful language processing requires completely automatic processing. I believe listening is the best way to achieve that, because it is the only skill where the speed of language processing is not determined by the user. A student can read and write at their own speed, and the speaker can speak as slowly as they want, whereas generally, the listener must process and comprehend the language at a speed determined by the speaker: usually quite quickly, at a rate of about three words a second. This forces the student to automate their language processing. This is the core of linguistic competence. As a corollary to this, good listening ability is also probably the best way to ascertain whether a second-language learner has successfully internalized the core of the linguistic system.

Of course, apart from requiring good levels of basic linguistic competence, listening comprehension also requires specific abilities not required by any of the

https://doi.org/10.1075/lllt.50.pre
© 2018 John Benjamins Publishing Company

other three skills. Firstly, since the information is conveyed through sound, successful listening requires knowledge of the sound system. The basis of this is the phonetic system. However, knowledge of basic phonetics is not enough. Since most people speak at a fast rate, the sounds are modified during rapid speech, resulting in considerable phonological modification (assimilation, elision, vowel reduction, etc.). The changes vary somewhat from one language to another, so learners must learn the phonology of the language. Intonation also carries considerable meaning, indicating not only the grammatical structure of the discourse, but also the relative salience of the information given. It may also indicate sarcasm, disbelief or provide clues about the speaker's true feelings. The intonation may even contradict the literal meaning, and at such times, the intonation is generally the truth the speaker is trying to convey.

The second specific ability relates to the fact that speech is usually a first draft, at least spontaneous informal speech, with the speaker figuring out what to say and how to say it, while they are saying it. Hence, speakers often use fillers, such as pauses, *ums* and *ahs*, as well as repetitions, as they take time to figure out what they want to say. And like all first drafts, there are usually a variety of infelicities and mistakes, which result in repairs, as speakers correct themselves, or think of a better way to convey what they want to say. Listeners must develop the ability to identify these fillers and other infelicities, and separate them from meaningful text. Of course, in some situations, especially more formal situations, the speaker may have time to prepare in advance, but such situations are not typical of most spoken interaction.

The third specific ability relates to the nature of the linguistic system itself. Spoken language tends to be far more colloquial, and less formal than written language. The vocabulary, grammar and discourse structure usually differ from written norms, often quite considerably. People do not speak in sentences, but in short idea units (often called *tone groups*), which tend to be shorter, often grammatically truncated, and strung together more by conjunction, rather than the embedding common in written language. The vocabulary is often colloquial, with some words or expressions used only in speech, often with more transitory, modern or fashionable expressions, and with far more slang. Connections between ideas are often not made explicit, but require the listener to make assumptions based on the juxtaposition of ideas rather than the cohesive markers common in written language. There is also much more personal involvement in spoken language, with far less concern for strict truth, and much more poetic license.

All the abilities mentioned above are language skills, in which listeners are using their linguistic knowledge to understand the speaker. But listeners use far more than linguistic knowledge to understand spoken text: they use whatever other information they think is relevant. Listeners must develop the ability to integrate

this information into their linguistic processing. Context is usually very important. The most obvious context is what the speaker said earlier, as each section of text becomes the context for interpreting later sections. If a speaker begins talking about one thing, the listener can usually make assumptions about what is likely to follow, and what is unlikely to follow. If a listener misunderstands the first section, they may lack the crucial context to understand what follows. The situation in which the speech takes place also creates important context, since the topic, as well as the language used, will typically vary considerably from one situation to another, often in predictable ways. An important part of this is the visual information available to the listener: especially the speaker's expression and general demeanor. Not only knowledge of the context, but also general background knowledge is crucial in successful listening. Speakers usually do not make everything explicit, but assume that the listener shares certain knowledge, ideas or assumptions, and expect them to incorporate those into their understanding.

The result of all this is that there are often many indirect speech acts. Speakers tend not to say what they mean in the precise way that writers usually do, and the listener is left with the task of not only understanding literally what the speaker is saying, but also of understanding what the speaker means by speaking in that way, in that context. What speakers mean may be very different from what they say. It is the integration of these four different types of knowledge – linguistic knowledge, knowledge of the co-text, knowledge of the situation, and general background knowledge – that leads to successful comprehension. Testing such a complex construct is a challenging enterprise, but there are some obvious characteristics that listening tests must have. We discuss four of these.

The first obvious characteristic a listening test needs is to test knowledge of the sound system. Not only the phonemes spoken in citation form, but the pronunciation of rapid, informal speech, complete with all the typical phonological modifications, as well as understanding the intonation. Written texts read aloud, for example, especially when read somewhat slowly and clearly, do not usually assess the test takers ability to process real spoken texts. This applies particularly to intonation; natural intonation is especially difficult to replicate in a non-spontaneous situation.

Related to this is the question of dialect, or speech variety. Pronunciation may vary considerably from one group of users to another, depending on geographical location, social group, race, and so forth. Furthermore, there are often differences between members of the same group. In some cases, there is a standard speech variety that might be chosen. But competent listeners can and should be able to adapt to deviations from standard norms, and this ability is an important listening skill. We all know through our daily experience that we all find some speech varieties more difficult to understand than others. This applies even more to L2 listeners, and test

developers need to make decisions about whose pronunciation to use. We need to know more about how much difference dialectical variation makes, or when it does or doesn't make a difference. This is an important practical question.

The second characteristic a listening test needs is to test the ability to understand all the hesitations, false starts, repetitions, repairs and simple errors typical of fast, spontaneous speech. The texts used need to be spontaneous spoken texts.

The third characteristic is the need to test knowledge of the vocabulary, grammar and discourse patterns typical of spoken language.

The fourth characteristic that needs to be tested is the ability to incorporate non-linguistic information into the comprehension process. We have seen that successful listening typically requires not just language processing, but also requires the incorporation and assimilation of other relevant information into the comprehension process. This is an important skill. But of course, in second-language testing, some might question whether we should be testing the use of non-linguistic knowledge, by replicating realistic listening tasks, or whether we should restrict ourselves to testing the linguistic knowledge necessary to be a competent listener. Opinions may differ, but there is a strong feeling among test developers that listeners should at least know the context in which the speech takes place, and further, they should be able to see the speaker, so they can pick up the common visual clues speakers provide. This is clearly an area that needs further exploration.

Not only do tests need to assess the above four characteristics of real-world listening, but they must assess them in an appropriate manner. After all, skills which cannot be used in real time are only partially acquired, and may not be very useful in real-life listening situations. It is fairly common to hear L2 learners explain their failure to comprehend by complaining that they could have understood the speaker, but the speaker was just speaking too fast. Of course, the reality is that the listener was listening too slowly! Most second-language learners, and perhaps most language teachers too, under-estimate the speed at which we process speech. The ability to process language at normal speech rates (say about 180 words per minute) is crucial to listening success. This must be addressed by our listening tests.

Perhaps the most common way currently used to test L2 listening ability is to write a passage, along with a number of comprehension questions, and then make a recording of someone reading the passage (often with slow, clear pronunciation). Such tests are so common that many teachers will have seen nothing else. They are found in simple teacher-made tests, as well as in prestigious, high-stakes tests. Such tests typically do not use spontaneous speech, do not use realistic spoken language, do not require the integration of contextual information, and often do not require the listener to process the language at normal speech rates. Such tests obviously do not test many important listening skills, and hence lack construct validity. While a case may be made that beginning listeners cannot process fast, colloquial speech,

and need to start learning with slower easier texts, serious tests must start using construct-valid spoken texts.

It seems obvious that listening tests must use authentic spoken texts. However, that doesn't mean that the texts necessarily need to be taken from some real-world source. What it means is that texts should have the characteristics of real-world spoken texts. Written texts read aloud are simply inappropriate. But if test developers are creative, they can surely find ways to create spoken texts with the right features.

However, creating listening tests with realistic spoken texts is far more difficult than it seems. Good test items are not easy to make, especially selected-response items. There is a complex interaction between the information in the text, and the stem (i.e. the question), the key (i.e., the correct option) and the distractors (i.e., the incorrect options). The correct option needs to be indisputably correct, and the distractors need to be indisputably incorrect. But given the ambiguity of language, especially spoken language, this is not so easy. Often, despite strenuous efforts to re-work the stem or the various options, it is just not possible to get the items to work properly. Furthermore, test developers are under pressure to write as many items as they can on each text – after all, more items mean more information about the test takers, and the more information we get, the more reliable the score. Test developers generally find that if they make modifications to the text, even small modifications, they can often make their items work, or work much better, or get more items out of the passage. It is easy to modify written texts, but how can you modify spontaneous, spoken texts?

This poses a serious dilemma. On the one hand, many of the distinctive characteristics of speech are due to the spontaneous nature of spoken language. Very few people, even the most accomplished actors, can create realistic texts from scripted input (the speech of films may seem realistic to the normal viewer, but generally it is quite different from real-world, spontaneous speech). Therefore, realistic spoken language needs to be spontaneous – this is its major defining characteristic. However, on the other hand, efficient test development often requires that texts be manipulated in some way. The dilemma is: should test developers use authentic, pre-recorded texts, knowing they'll have trouble writing enough good items (or perhaps even any good items), or should they use pre-designed, non-spontaneous texts, because they can write much better test items. Or to put this in more technical terms: do they test the right construct poorly, or do they test a closely related construct well?

Of course, this problem arises because test developers are trying to separate listening from speaking, and test listening as a separate skill. Clearly this has advantages. Such listening tests can be administered to many students simultaneously, and can even be machine scored. This reduces cost and increases convenience considerably. But group administration requires pre-recorded listening texts, and easily

scorable tasks, with all the problems that follow. In the real world, listening usually takes place in an interaction with a speaker, or with multiple speakers (although radio, and digital media have changed that to some extent). So why not test listening as part of a spoken interaction with a speaker? There are obvious problems, to say nothing of expense and inconvenience, but it is easy to imagine how we could assess important listening skills by replicating real-world spoken interaction. Clearly this is an area for more research.

The issues in testing listening are very complex, and this current volume attempts to address many of them. There are four main themes, all of which are of practical importance to our understanding of how to test second-language listening ability:

1. The use of authentic spoken texts,
2. The effects of different types, and different strengths, of accent,
3. The use of audio-visual texts on L2 listening tests,
4. Assessing listening as part of an interactive speaking/listening task.

Listening comprehension is an important component of second language ability, and yet testing listening in a reliable and valid manner is still a challenge for our field. Not only teachers, but also large, well-funded test publishers do a fairly poor job of this. The theory is pretty well understood, but as a field, we just do not know enough to implement that in a practical manner. Research into these four topics will add considerably to our understanding of second language listening, but more importantly, it should provide useful guidance on the practical problem of how to better test this challenging construct. We need much better listening tests. The authors are to be commended for their efforts. I hope their work will encourage others to do the same.

Gary Buck
Monterey, CA, 2017