21

INDIVIDUAL DIFFERENCE FACTORS FOR SECOND LANGUAGE LISTENING

Franz Holzknecht and Tineke Brunfaut

Background

For many years, listening was underrepresented in second language research, despite often being regarded as “the most fundamental skill of all” (Buck, 2018) for first language (L1) and second language (L2) acquisition. Although the past two decades have fortunately seen a considerable increase in studies, the field is still far from a fully comprehensive understanding of this crucial language skill. One reason for the limited insights into listening is its complex and covert nature compared to other modalities such as speaking and writing, which can be observed directly, and reading, where insights might be gained through methods such as eye tracking.

The most commonly cited models of listening comprehension in the L2 literature are those proposed by Vandergrift and Goh (2012), Field (2013), and Rost (2016). Although these three models differ slightly in how listening is conceptualized, they all agree that listening comprehension involves both bottom–up (neurological and linguistic) and top–down (semantic and pragmatic) processes. While neurological processing underlies all other processes and involves the physiological activities of the ear, and thus hearing (Rost, 2016), the linguistic processes can be split up into input decoding (i.e., discerning whether an incoming audio signal is speech), lexical search (i.e., recognizing individual words), and syntactic parsing (i.e., organizing individual words into a syntactic pattern). These bottom–up processes are in turn influenced in a top–down fashion by semantic and pragmatic processes (i.e., through prior knowledge of words and phrases, as well as world knowledge) and inform the higher-level processes of meaning construction (i.e., situating sentences in the context of the speech situation), and discourse construction (i.e., identifying the meaning of the discourse as a whole). However, none of the processes can be observed directly but have to be inferred from performances on listening tasks (Rost, 2016). What adds to the challenge is that the processes act in a parallel, interactive manner rather than as serial patterns, while also being constantly updated as speech continues to come in (Field, 2013; Vandergrift & Goh, 2012).

While many listening researchers use introspective techniques to learn how L2 listeners make sense of spoken language, another way to further the field’s understanding of L2 listening is to examine how individual differences (IDs) between learners affect listening development and performance. In this chapter, IDs are conceptualized as learners’ unique characteristics, traits, and dispositions influencing their listening processes and listening outcomes. Although investigating such ID factors does not offer direct insights into the nature of the skill itself, it can help us understand which learner variables contribute to successful listening. This in turn can inform language teaching, language testing, and educational policy about effective methods to teach and assess L2
Franz Holzknecht and Tineke Brunfaut

listening. For example, if it was found that anxiety as an ID factor negatively impacts L2 listening performance, steps could be taken to reduce listeners’ anxiety levels in the language classroom and in language tests.

This chapter aims to discuss the current state of research on ID factors in L2 listening. To this end, and guided by the selection of IDs covered in this handbook, a literature search was carried out, starting off from well-known outlets in the L2 listening literature (1–3) and supplemented with a comprehensive database search (4):

3) Monographs and edited volumes on L2 listening, and publications by highly cited and well-established L2 listening scholars, e.g., John Field, Christine C. M. Goh, Michael Rost, and Larry Vandergrift, to name a few.
4) An online search, including the 272 databases subscribed to by the authors’ institutions, using the list of IDs covered in this handbook in combination with “listening” as search terms.

Each of the relevant studies identified was then classified according to one or more of the terminological categories discussed in Parts 2–5 of this handbook. While aiming for a good representation of work on IDs in L2 listening in this chapter, it should be noted that the length restrictions of a publication of this nature prevent the inclusion and coverage of all relevant existing research.

The chapter is divided into four main sections. Following this Background section, the existing research base on ID factors in L2 listening is reviewed (see the Research section below), with the first part describing the empirical findings of relevant research identified in the literature search (see Evidence). In the second part, entitled Data Elicitation, the main research methods used in studies investigating these difference factors are outlined. The third main section of this chapter, Practical Applications, examines the research findings in terms of their implications for practitioners such as language teachers, language testers, and policy makers. The chapter concludes with Future Directions, identifying important research gaps in this topic area and offering recommendations for further research.

Research Evidence

Focusing on the IDs represented by the different chapters of this handbook, this section reviews studies that specifically investigated those IDs in the context of L2 listening. Parallel to the structure of the handbook, the research is discussed according to cognitive differences (working memory, learning strategies, and metacognition), conative differences (motivation and willingness to listen), affective differences (anxiety, self-efficacy, and learner beliefs), and sociocultural and demographic differences (identity and age). For construct definitions and models of the IDs themselves, we refer the reader to the relevant ID chapters in this handbook.

Cognitive Differences

Working Memory

As discussed in Chapter 3 of this handbook, working memory (WM) differences between learners have been theorized, and empirically demonstrated, to impact the acquisition of a second language.
The role of working memory capacity has also been investigated in relation to L2 listening comprehension, mostly in combination with other factors. One of the first studies on this particular topic is by Kormos and Sáfár (2008), who examined the relationship between WM and listening scores for 121 Hungarian secondary-school students taking the Cambridge First Certificate Exam (FCE). The study examined both phonological short-term memory (PSTM), which only taps the storage function of working memory, and complex working memory, which involves both the storage and executive functions of working memory. While PSTM and listening performance did not show a significant correlation, a moderate correlation was found between complex WM capacity and FCE listening scores. That is, students with high scores on the complex WM test also tended to achieve high scores on the listening exam. Similar associations between complex WM capacity and L2 listening performance (and somewhat weaker ones between PSTM and L2 listening) are also reported by Brunfaut and Révész (2015), who looked into the role of WM (among other variables) in the English listening test performances of 47 Chinese L1 speakers. Based on their findings, the authors suggested that “individual differences in WM predict L2 processing abilities” (p. 161), but they highlighted that the strength of the correlation depends on task type and subskills tested. The authors argued that WM seems to play a role in listening comprehension only in listening items with high reading load (i.e., a multiple-choice stem formulated as a question and three sentence-length answer options) and which targeted detailed listening comprehension processing to a larger extent. For shorter listening items with very limited reading load and targeting more global listening skills (i.e., overall rather than detailed understanding), no correlation between WM and listening performance was found.

Other research revealed similarly mixed results. For instance, in a study involving 121 native and 113 non-native speakers of Dutch, Andringa et al. (2012) investigated, among other factors, the relationship between WM and L1 and L2 listening success in a multisample structural equation modeling (SEM) analysis. The authors found only weak correlations between scores on different WM tests and an advanced test of Dutch as an L2. Overall, WM did not explain any unique variance (beyond language knowledge and IQ) in listening performance. Similar findings are reported by Vandergrift and Baker (2015), who examined the role of WM in L2 listening comprehension with a sample of 157 Canadian learners of French, split into three cohorts. The authors found a significant, moderate-sized correlation between WM capacity (backward digit recall and nonword list recall) and French listening scores (mostly detailed comprehension) for one of the three cohorts in the sample (47 students). However, no relationship was found for the remaining two cohorts (110 students), who were thought to have similar L2 listening proficiency.

Thus, overall, the limited research available suggests that WM capacity predicts differences in L2 listening comprehension only to some extent. Detailed comprehension and listening tasks with more reading needed in the comprehension questions (i.e., stems and answer options) may be more taxing on WM than global processing (i.e., understanding the overall meaning of a listening text rather than the specific details within the text) and shorter tasks.

Learning Strategies

While Chapter 5 provides a solid review of language learning strategies, the focus here is specifically on strategy use for L2 listening. Two strands of research prevail in the latter area: 1) studies investigating learners’ listening strategy use and the relationship with their L2 listening proficiency; and b) studies looking into the effectiveness of different L2 listening strategy instructional methods. An example of the first type of studies is Shang (2008), who explored the listening strategy use of 97 English as a foreign language (EFL) university students in Taiwan in relation to their EFL listening proficiency level, as determined by a simulated Test of English as a Foreign Language (TOEFL) listening module. Shang focused on three different linguistic patterns in the listening passages. While all learners drew on memory strategies, low-level learners reported relying more
heavily on memory strategies, especially to comprehend negative expressions in the listening passages. Advanced L2 listeners, on the other hand, seemed to employ much more of a combination of different listening strategies, especially to understand contrary-to-fact statements in the listening passages. The study thus shows that more proficient listeners not only use a greater variety of listening strategies, but they also seem to use these strategies more effectively. However, it is less clear whether strategy use constitutes a learner ID or reflects listening proficiency level as such.

Other research also studied the relationship between learners’ listening strategy use and their L2 listening proficiency. Kök (2018), for example, investigated cognitive, metacognitive, and socio-affective strategies of 44 EFL Turkish university students at upper-intermediate and intermediate levels by means of a Listening Comprehension Strategy Use Inventory. Both cognitive and metacognitive listening strategy use was found to correlate strongly with L2 listening, as measured by the International English Language Testing System (IELTS). Kök also found that those with higher metacognitive strategy use had higher L2 listening proficiency as opposed to those with higher employment of other strategy types. Nix (2021), in a multicomponent SEM study with 618 EFL university students, found that bottom–up strategy use resulted in lower L2 listening achievement, whereas the joint combination of bottom–up and top–down processing strategies had a positive effect on L2 listening. Graham et al. (2011) similarly arrived at a nuanced picture of strategy use and L2 listening. They found that learners with different French–L2 listening development patterns, while using similar types of strategies, differed in the proportion of reliance on specific strategies in their combination patterns of strategies and in their compensation strategy patterns.

One particular type of strategy, which becomes increasingly relevant as more listening events are taking place in a digital online environment, is the self-regulation of audio recordings (i.e., being able to rewind, pause, fast-forward, and change volume). Ozcelik, Van den Branden, and Van Steendam (2019) investigated how 28 A1-level English learners in Turkey exploited self-control of audio recordings (rewind, pause, and fast-forward) during peer interactive listening tasks. They found that the learners took advantage of the self-regulation option in 66% of problems experienced, and most often did so to verify their comprehension of the listening input, identify task-relevant information, or adjust for issues experienced with delivery speed. The authors argued that the self-regulation had helped keep down learners’ anxiety levels during the demanding tasks.

The second type of studies, i.e., exploring the effects of strategy instruction on listening comprehension, showed mixed findings. Yeldham (2016) reported a quasi-experimental study with 67 lower-intermediate EFL students at a Taiwanese university exploring the impact of two different instructional methods on L2 listening proficiency (as measured by the General English Proficiency Test [GEPT]). Both groups underwent a 22-hour course over two semesters, with one group explicitly taught and practicing a set of cognitive and metacognitive strategies for listening (“strategies group”), and the other group additionally taught and practicing a set of bottom–up, lower-level listening decoding skills (“interactive group”). Comparative statistics on the pre-test–post-test design did not reveal a significant difference between the two groups. However, within the strategies group, learners’ EFL listening proficiency significantly improved over time, whereas this was not the case within the interactive group, and larger effect sizes were found for certain aspects of strategies growth within the strategies group. The author concluded that “for lower-intermediate level listeners, it is better to focus more on developing their listening strategies than to provide them with a balanced interactive approach” (Yeldham, 2016, p. 394). In another, small-scale intervention study of Japanese-as-a-foreign-language learners at an Australian university (Seo, 2005), one group of learners received modeling and practice in the strategies identifying key terms, elaborating, and inferring while listening, during a two-semester course. This intervention group was shown to outperform a control group in terms of L2 listening by the end of the course. A further example of a study on the effectiveness of listening strategy instruction, which identified a positive impact on L2 listening, is Graham and Macaro (2008). Roussel et al. (2019) similarly identified the positive effects of low-level (categorizing and segmenting individual words and phrases)
and high-level (using world knowledge and context knowledge) process-based training on 108 L2 listeners’ comprehension levels. However, they argued that the effect of the strategy instruction importantly depended on the L2 listeners’ proficiency levels at the start of the treatment. More specifically, a couple of studies, such as Cross (2009) and Park (2010), did not find significant L2 listening gains of the strategy use instruction explored in their respective studies.

A highly influential body of experimental studies on metacognitive instruction in the context of L2 listening is that by Larry Vandergrift and colleagues. For example, Vandergrift and Tafaghodtari (2010) explored the effectiveness of a program that guided French-L2 listeners through the metacognitive processes of prediction, planning, monitoring, evaluating, and problem solving. They found that the experimental group (n = 59) outperformed the control group (n = 47)—who had not received any such guidance during the listening course—in listening comprehension achievement (and also in listening metacognitive awareness). Similar metacognitive instruction studies by Cross (2011), Rahimirad and Shams (2014), and Mahdavi and Miri (2019) likewise identified positive effects on EFL listening comprehension (and the latter two studies also on metacognitive awareness), although in the case of Cross’ study this was primarily associated with the less-skilled listeners in the study (with little or no listening gains for more-skilled listeners).

It thus seems that 1) differences in listening strategy use are associated with differences in L2 listening proficiency level (i.e. more proficient listeners seem to use a greater variety of listening strategies more effectively), 2) certain strategy instruction methods may be more effective than others at lower-proficiency levels, and 3) more proficient L2 listeners may not benefit from strategy instruction as much as less proficient L2 listeners. It is less clear from the existing research, however, whether listening strategy use constitutes a learner ID or is a reflection of listening proficiency level as such (i.e., whether strategies are a cause or an effect).

**Metacognition**

For a domain-general definition of metacognition, see Chapter 6 of this volume. In the context of L2 listening more specifically, Vandergrift and Goh’s (2012) work has been instrumental, and defines metacognition as comprising: 1) metacognitive experiences, i.e., feelings or thoughts around cognitive aspects of listening; 2) metacognitive knowledge, i.e., knowledge and beliefs about variables related to the person, task, and strategies involved that impact cognitive aspects of listening; and 3) strategy use, i.e., the employment of strategies to manage and regulate listening. The latter has been discussed in the preceding section; here, the focus will be on metacognitive experiences and metacognitive knowledge.

A substantial contribution to L2 listening research in this area has been the development and validation of the Metacognitive Awareness Listening Questionnaire (MALQ) by Vandergrift et al. (2006). Following factor analyses on data from 966 and 512 L2 listeners from a range of backgrounds and proficiency levels, the final version of the questionnaire consisted of 21 items assessing the constructs of problem solving, planning and evaluation, mental translation, person knowledge, and directed attention. Vandergrift et al.’s (2006) study identified a significant relationship between L2 listeners’ metacognitive awareness and their listening comprehension levels. Since its publication, the MALQ has been adopted in studies on L2 listeners’ metacognitive awareness in a range of contexts, typically concluding a positive relationship with L2 listening (for an overview, see Goh & Hu, 2014). Goh and Hu (2014), for example, used the MALQ and went into greater depth than most other studies by looking into how the different components of metacognitive awareness associate with L2 listening. For a group of 113 Chinese English-as-a-second-language (ESL) learners at a university in Singapore, they found that metacognitive awareness explained 22% of the variance in L2 listening performance, with the strategies of directed attention, problem solving, and confidence relating significantly with listening. They also identified variation between learners in different components of metacognitive awareness. However, in a recent study by Nejad
and Farvardin (2019) with 100 lower-intermediate EFL learners in Iran, only a weak correlation was found between metacognitive awareness and L2 listening. Nejad and Farvardin additionally looked into aural vocabulary knowledge, which proved to be a much stronger correlate with listening.

Conative Differences

Motivation

Motivation has been found to be a highly influential ID for L2 learning (for an overview, see Chapter 7). Few studies, however, have specifically zoomed in on motivation and L2 listening development. An innovative series of large-scale studies on the English language capability of learners in Spain (Madrid), Colombia (Bogotá), and Sri Lanka (countrywide) (Shepherd & Ainsworth, 2017, 2018a, 2018b) investigated the relationship between English listening proficiency (measured with the Aptis listening test) and English language learning motivation (measured with scales operationalizing ideal L2 self, ought-to L2 self, language learning experience, instrumentality, international orientation, English self-concept, parental encouragement, and motivated learning behavior). Several motivational variables correlated highly with Madrid-based L2 listeners’ proficiency, and moderately with Bogotá- and Sri Lanka-based L2 listeners’ proficiency. In all three settings, the following motivational variables showed comparatively larger positive correlations with L2 listening (although with some variations in strength between settings): learners’ self-beliefs of the ability to study English (English self-concept); learners’ image of themselves as proficient English speakers (ideal L2 self); learners’ interest in learning English (language learning experience); learners’ perceptions of the usefulness of English for their future (instrumentality); and learners’ interest in English for global communication (international orientation; in Madrid and Sri Lanka only).

Another study that investigated the relationship between motivation and L2 listening comprehension is Vandergrift (2005), who found for French–L2 young learners that higher levels of amotivation (not seeing the relationship between one’s actions and their consequences) were associated with lower L2 listening performance, but that intrinsic and extrinsic motivation did not correlate with L2 listening. It should be noted that this line of research on the impact of motivation on L2 listening usually focuses on L2 listening motivation in general and not specifically motivation for L2 listening development.

A study that more narrowly defined motivation for L2 listening, and specifically for the testing of L2 listening, is Xu (2017). It was found for a group of 560 Chinese undergraduates that their test-taking motivation for the CET-4 listening test (i.e., their expectancy beliefs about and interest in this listening test, but not their importance perceptions of the test) impacted their listening scores. Noteworthy, while these motivation factors had a direct effect on students’ listening test scores, Xu also found that this relationship was at the same time mediated by the students’ listening metacognitive awareness (measured with the MALQ).

As part of the research on motivation, explorations of attitude are often included. Most studies have focused on attitudes towards L2 learning in general (see e.g., Gardner, 2010), while very few have considered attitudes towards L2 listening specifically. Takkaç Tulgar (2021) is an exception, which examined the listening-related attitudes of L2-Turkish and L2-English learners in Turkey. The study found that both groups held positive attitudes towards listening in their L2 and developing their L2 listening skills, despite differences between the two languages in the amount and type of listening opportunities in the study’s setting. It would be relevant to complement this type of research on the nature of L2 listeners’ attitudes with studies that explore the relationship between attitude and learners’ L2 listening ability. More importantly, however, it is often unclear how attitude has been defined in L2 listening studies, including in Takkaç Tulgar’s (2021) study where the interview questions suggest it may have been conceptualized as learner beliefs. It is therefore imperative that future research in this area prioritizes substantiating the construct of attitude itself,
and establishes how it can then be operationalized in a meaningful way, before empirically investigating its potential effect on L2 listening.

**Willingness to Listen**

A considerable body of literature exists on willingness to communicate (WTC) in a second language, i.e., L2 learners’ readiness to enter into communication with others (for an overview, see Chapter 10). The focus in that tradition has been on learners’ willingness to produce language—primarily speaking in the L2—although some studies investigated writing in the L2. In L1 research, Roberts and Vinson (1998), however, explored people’s willingness to listen (WTL). They argued that while willingness to speak or write typically is an intentional act, listening involves a mix of conscious and nonconscious engagement with, and intensity adjustments to, aural channels. Listeners’ decisions on the level of effectiveness with which they listen might depend on, for example, their interest to learn or their judgment of how valuable the aural information is likely to be; in other words, their motivation to listen and their monitoring practices. Using a self-developed WTL scale, Roberts and Vinson found for a group of 94 US university students that WTL correlated positively with the students’ self-perceived communication skills but negatively with their anxiety (both listener- and speaker-based apprehension) and dogmatism. The researchers also concluded that WTL is distinct from WTC.

Our literature search did not reveal any research on WTL in an L2. Exploring this topic, therefore, seems a guarantee for novel research in second language acquisition (SLA)—both in terms of investigating L2 learners’ WTL in its own right and in relation to L2 listening proficiency.

**Affective Differences**

**Anxiety**

It is well documented that anxiety can impede language learning in general (Horwitz, 2010; see also Chapter 11 of the current volume), and the role of anxiety in L2 listening comprehension, in particular, has also been investigated in several studies. For example, Brunfaut and Révész (2015) included listening anxiety as one factor in their investigation of the role of L2 listener characteristics in L2 listening performance. By correlating responses to a listening anxiety questionnaire with listening test scores, the authors were able to show that anxiety was higher in test takers with lower L2 listening performance. In other studies, results were similar in that listening anxiety, as indicated through questionnaire responses, was negatively correlated with listening comprehension scores (Elkhafaifi, 2005; Golchi, 2012; Kim, 2000; Kimura, 2017; Xu, 2017; Yang, 2010; Zhang, 2013). In addition, Kimura (2017) showed that listening anxiety may consist of two distinct but related constructs—self-focused apprehension and task-focused apprehension—and that both are enhanced through high levels of general social anxiety.

In contrast, however, in a study that specifically focused on the testing of L2 listening and explored the role of general test anxiety rather than listening anxiety or listening test anxiety, In’nami (2006) did not find a significant relationship between test anxiety and listening test performance. The study involved 79 EFL university students in Japan who completed two TOEFL listening tasks and two test-anxiety questionnaires representing general test worry, test-irrelevant thinking, and emotion. The fact that the listening test had low stakes for the students, and thus may have been less anxiety-inducing, might (at least partly) explain the findings.

Thus, while it is not so clear how and whether test anxiety relates to L2 listening test performance, research findings seem to show that listening anxiety and weaker L2 listening performance are related, although correlational studies as such do not clarify which variable affects the other (or both do together). A number of other studies, however, have shown that anxiety (listening anxiety and interestingly also test anxiety) can be reduced to some extent, with positive effects on listening...
performance. Arnold (2000), for instance, showed that ESL students instructed in visualization-relaxation techniques were less anxious and performed better on listening tasks than a control group. Similarly, Li, Wu, and Lin (2019) report that brainstorming techniques prior to listening tests can help reduce test anxiety and foster comprehension, while Holzknecht’s (2019) findings demonstrate that test takers are significantly less anxious when completing listening tests in double play (i.e., the listening texts are played twice) compared to single play. Holzknecht’s (2019) results thus partly confirm findings by Ozcelik, Van den Branden, and Van Steendam (2019), who report that self-control of audio recordings helped to reduce listeners’ anxiety, as discussed above.

In sum, although listening anxiety seems to be a clear ID factor in L2 listening, it can also be mitigated through different techniques.

Self-Efficacy

Self-efficacy relates to learners’ task-specific beliefs in their abilities (see also Chapter 13) and is speculated to be especially important for L2 listening. However, the only study available which directly investigated its role in L2 listening seems to be Graham and Macaro’s (2008) research on the effects of strategy instruction on the listening proficiency of 68 learners of French. The authors report that strategy instruction improved learners’ listening performance and also enhanced their self-efficacy. While these findings are promising, more research on the role of self-efficacy as a potential ID factor in L2 listening is clearly necessary.

Learner Beliefs

For a review on learner beliefs as an ID in general, we direct the reader to Chapter 14. With specific reference to L2 listening, there appear to be only two studies on what learners believe about L2 listening and how this influences their listening ability. In the first of these (Nix & Tseng, 2014), the authors studied young Chinese learners’ beliefs about English listening ability by means of a questionnaire. They categorized learner beliefs according to “axiomatic beliefs”—what learners think listening ability to be (e.g., “English listening ability is an important part of overall English ability”)—and “praxis beliefs”—how learners think listening ability should be applied in practical terms (e.g., “I think I should listen to native speakers of English as much as possible”). In the second study (Nix, 2021), a follow-up on the first, the author showed that affirmative “axiomatic beliefs” weakly correlated with higher listening ability as expressed through test scores, but affirmative “praxis beliefs” did not associate with higher listening scores. While these results are interesting, more research is needed on the relationship between learners’ beliefs about listening and their L2 listening ability.

Sociocultural and Demographic Differences

As discussed in Chapters 15 and 16, identity- and age-related factors constitute key sociocultural and demographic differences in language learners. In what follows, we specifically look into these IDs for L2 listening.

Identity: Ideology, Race/Ethnicity, Gender/Sexuality

The influence of test takers’ identity on their L2 listening performance has received relatively little attention from researchers. An early study is Markham and Latham (1987), who examined the relationship between religious background knowledge and L2 listening comprehension with 65 multinational adult learners of English. The authors found that when listening to a passage about a particular religion, participants pertaining to that religion were able to recall significantly more major idea units (with fewer distortions) than participants of a different religion as well as non-religious participants. The study thus highlights the important role of background knowledge for L2 listening comprehension.
Most research on the role of identity in L2 listening comprehension appears to have explored gender as an ID factor. For example, Park (2008) looked at differential item functioning (DIF) across gender in L2 listening in the context of the 2003 Korea College Scholastic Ability Test—a test that was not piloted prior to operational use. Park’s analysis of 40,000 Korean test takers’ responses revealed DIF for 13 of 17 English listening items, with seven items functioning in favor of females and six items in favor of males. Park identified four variables which were causing DIF to varying degrees: test content, use of pictures, dialogue or monologue, and question type. Based on these findings, the author argued that such variables should be considered when designing L2 listening tests. The study also highlights the importance of field trialing items before operational testing.

In two other studies, gender differences in L2 listening were investigated in combination with the effects of nationality or ethnicity. Aryadoust (2012) examined DIF across gender and nationality in the IELTS listening module with 209 multinational university students. While none of the items displayed DIF across nationality, several items functioned in favor of either females or males, but DIF across gender was also always related to candidate proficiency. Based on his findings, Aryadoust argued that lower-proficiency male listeners may be more prone to guessing than female listeners. In a similar study, Seo et al. (2016) did not find DIF across gender in a sample of 710 Asian and 523 European students taking a “large-scale English language proficiency test” (p. 52). However, the authors identified two items that displayed weak differential functioning between students of Asian and European heritage. The authors suggested the further investigation of such items prior to operational use to avoid introducing bias into test scores.

In summary, the existing body of research findings suggests that L2 listeners’ identity characteristics are only a factor in their listening performance differences if listening assessment instruments are not carefully developed and pre-tested. However, studies on identity variables other than gender are sparse.

Age

Most research on the role of learners’ sociocultural and demographic background in L2 listening has focused on age effects. Banerjee and Papageorgiou (2016), for instance, investigated DIF across three different age groups of L1 Spanish listeners (<17 years, 17–27 years, >27 years) in the context of the Michigan English Test. The authors focused on listening test items within the occupational domain, hypothesizing that older listeners with more workplace knowledge would perform better on these than younger listeners with less workplace knowledge. However, contrary to the hypothesis, the findings indicated that “items in the occupational domain did not consistently demonstrate DIF with younger test takers” (Banerjee and Papageorgiou, 2016, p. 19). Similar results were reported in the studies by Aryadoust (2012) and Seo et al. (2016) outlined above. Aryadoust (2012) investigated L2 listeners aged 16 to 49 and Seo et al. (2016) examined L2 listeners in Grades 9 to 12 (presumably aged approximately 15 to 18, although the authors do not mention the exact age). Neither of these two studies found DIF across the different age groups. Finally, Harley (2000) looked at differences in L2 listening strategies in a small group of L1 Polish listeners. Her findings “did not provide any evidence of age-related differences in L2 listening strategies” (p. 775), thus mirroring the results of the other studies.

Slightly different results were reported by Geranpayeh and Kunnan (2007) in their study on the Cambridge Certificate of Advanced English listening test. The authors looked at DIF across three different age groups and identified six out of 32 items as functioning differentially for the two focal groups (≤17 years and ≥23 years) in relation to the target age group (18 to 22 years). However, no clear patterns emerged in expert judgments as to why these items displayed DIF, leading the authors to conclude that the listening test was “probably not biased against the test taker age groups included in this study” (p. 207).
Thus, overall, research findings across five different studies largely agree that age does not seem to be a factor in L2 listening comprehension. However, it should be noted that all the studies involved large-scale language tests which were professionally developed and pre-tested. Similar to other sociocultural and demographic background knowledge variables, age-related knowledge still has the potential to introduce bias into listening test scores in language tests that are not carefully designed and piloted.

Data Elicitation

The previous section provided an overview of empirically based insights into the role of IDs in L2 listening, to date. The present section considers methodological aspects of the current research base, specifically the populations explored and the type of research instruments and data collection methods used.

As can be seen from the above descriptions of existing studies, ID factors in L2 listening have been explored in a number of different contexts, for a number of different target languages, and for learners from a number of different backgrounds. Overall, however, it seems that, so far, most insights into this topic have been gained for (young) adults listening in Germanic and Romance L2s (see the details on the target L2s and participant groups of the studies in the Evidence section above, e.g., ESL, French-L2, Dutch-L2, and university students). There are therefore plenty of opportunities to extend the current research base to younger L2 listeners (as well as older adults) and to a much wider range of L2s.

Our review also identified the type of data elicitation instruments that are used in IDs research in L2 listening. These include, on the one hand, measures to establish participants’ L2 listening ability, and on the other hand, instruments to measure the ID. With respect to L2 listening comprehension measures, in many studies these involve a large-scale, standardized listening proficiency test. Utilizing good-quality listening assessment instruments is crucial to study ID variables and we strongly encourage researchers to use tests that have been professionally developed by trained item writers and pre-tested on a suitable sample. Useful listening proficiency tests will normally include multiple listening inputs representing a range of text characteristics (e.g., genre, linguistic features, topic, and abstractness), a variety of task/item types (e.g., short-answer questions, multiple-choice, and table completion), and a decent number of items targeting several different listening subskills/processes (e.g., gist, main ideas, supporting details, and specific information) in order to ensure comprehensive construct representation. To capture differences in listening proficiency, there will need to be a good mix of item difficulties—in relation to the (anticipated) proficiency level(s) of the learner sample. To avoid construct-irrelevant factors, the reading, writing, and memorization demands to demonstrate listening comprehension through the items should be kept to a minimum, and items should not be answerable based on background knowledge. It needs to be kept in mind, however, that even when using existing standardized tests, these still need to be evaluated in terms of their specific intended use in a particular study and their appropriateness for the participant population. For guidance on selecting, adapting, or developing L2 listening tests, the reader is referred to Batty (2021).

With respect to the ID measures, research into ID factors in L2 listening typically involves self-report instruments (e.g., anxiety questionnaires), with the specific instrument used depending on the ID under focus. For example, when studying IDs such as listening strategies, metacognitive awareness, anxiety, learner beliefs, motivation, or WTL, most researchers have relied on questionnaires and scales. While some of these have been developed and validated for use in L2 listening research specifically (e.g., the MALQ by Vandergrift et al., 2006), many studies have adopted more general questionnaires, such as the Foreign Language Classroom Anxiety Scale developed by Horwitz et al. (1986). However, while FL classroom anxiety and listening anxiety are related with one another and with listening performance, research by Elkhafaifi (2005) has shown that these...
two types of anxiety can be empirically distinguished and may therefore be separate constructs. One reason for this may be that FL classroom anxiety is task specific, whereas listening anxiety is skill specific. Similarly, language learning motivation questionnaires are unlikely to capture motivation fully or specifically for L2 listening development, and thus there may be a need for skill-specific motivation questionnaires.

Data on learners’ WM has, alternatively, been collected with WM tests. Several (validated) test batteries are available for different populations, and different tasks tap into different components of WM (see Chapter 3, but also Wen et al., 2021). Key considerations in WM and L2 listening research, however, are the evaluation of 1) which WM components are associated with listening and subskills of listening (e.g., top–down vs. bottom–up listening), and 2) whether the role of working memory in listening is domain specific, for example, whether working memory measured via listening span tests shows stronger correlations with listening than reading span tests.

Qualitative methods have also been used to investigate IDs in L2 listening, e.g., verbal protocols (mostly in studies on strategies and metacognition) and individual/focus group interviews (in research on attitude). As these methods generate rich and insightful data when employed with care, there is potential for use with other ID factors too. Furthermore, as demonstrated by Holzknecht’s (2019) L2 listening study, qualitative approaches are highly valuable for triangulation with quantitative methods like questionnaires. Other qualitative methods such as diaries or experience sampling, used for example in WTC studies, are also worth considering.

For useful methodological discussions and recommendations on how to measure various IDs, and what instruments are available, the reader is referred to other chapters in this volume, as well as to the various chapters in Winke and Brunfaut (2021; Part II: Building instruments for SLA research & Part III: Measuring individual differences).

**Practical Applications**

Although studies on many of the themes discussed in this chapter are still sparse, our review of the current state of research on ID factors in L2 listening has revealed a number of areas relevant for practitioners. Firstly, similar to other areas of L2 learning, motivation seems to be an influential ID for L2 listening, as for example shown in the research reported by Shepherd and Ainsworth (2017, 2018a, 2018b) discussed above. It is thus important that curriculum designers and language teachers strive to develop and conduct L2 listening activities that are engaging for their students and that can raise students’ interest in their own listening development. It will also be worthwhile to raise students’ awareness of the usefulness of L2 listening skills in order to make the instrumental role of listening proficiency explicit. Furthermore, as shown by Xu (2017), test-taking motivation affects L2 listening test performance, and thus language teachers should aim to instill positive interest in learners towards listening assessments, which is likely to be enabled by good-quality and level-appropriate listening tests, as well as explicit communication around the purposes of such tests (e.g., confirming learning achievements, or diagnosing learner’s strengths and weaknesses to provide additional support). It will also be helpful to assist learners in developing accurate insights into their own listening competence and accurate self-evaluations of task performance in order to enable positive expectancies and self-concepts as learners of L2 listening. Useful resources on developing learners’ motivation and subsequent engagement are Mercer and Dörnyei (2020) and Dörnyei and Ushioda (2021).

Another ID factor in L2 listening with practical implications is anxiety. Studies on this topic have consistently revealed a negative relationship between students’ L2 listening anxiety and their L2 listening performance (see above). While the intervention studies discussed above indicate that lowering learners’ anxiety levels improves their L2 listening performance, it remains to be established whether, in addition, lower-performing students are also more anxious in general and thus whether lower listening proficiency also systematically causes greater anxiety. Regardless of the
exact causal relationship between these variables, it is clear that reducing students’ L2 listening anxiety can still be beneficial, and researchers have begun to identify methods that can help reduce anxiety in relation to L2 listening. It has been shown that visualization-relaxation techniques (Arnold, 2000), brainstorming techniques (Li et al., 2019), playing the listening text twice instead of once only (Holzknecht, 2019), or giving learners full control over playing the audio recordings (Ozcelik et al., 2019) can reduce learners’ L2 listening/test anxiety. Whenever feasible, language teachers and testers should thus try to incorporate these methods in their lessons and assessments. At the same time, researchers should continue to explore different ways to reduce students’ anxiety related to L2 listening.

Language teachers, language testers, and curriculum developers should also be considerate of potential sociocultural and demographic difference factors for L2 listening. It is promising that most research in this area has demonstrated that factors such as personal identity, nationality or ethnicity, gender, or age do not seem to play a role for IDs in L2 listening, when listening exercises or tests are carefully developed (e.g., as is the case in many large-scale, standardized language tests). However, when listening assessment instruments are not professionally designed and pre-tested, sociocultural and demographic factors can strongly influence L2 listening success (e.g., Park, 2008). Practitioners involved in L2 listening should therefore ensure that they choose topics that are equally accessible to all students, regardless of the students’ sociocultural or demographic backgrounds. They should familiarize themselves with principles and practices of good listening task and test design, and they will find it useful to consult resources such as Batty (2021), Buck (2001), and Green (2017).

Apart from motivation, anxiety, and sociocultural and demographic variables, there are also other practically relevant areas that could potentially constitute ID factors for L2 listening; however, research has been often too sparse to be able to offer clear guidance for practitioners. For example, as discussed earlier, there are some indications that learners with a higher working memory capacity might also be more successful in L2 listening, particularly for tasks with high reading loads and focusing more on detailed comprehension (Brunfaut & Révész, 2015). While our literature review did not uncover experimental studies on the effect of WM training on L2 listening comprehension, L1 research with young children at risk for learning difficulties suggests that working memory training may be beneficial for L1 listening comprehension (Peng & Fuchs, 2017). It thus seems worthwhile to explore the nature and effectiveness of WM training in an L2 context as well.

The current research strongly suggests that listening strategy use and metacognitive awareness are associated with differences in L2 listening proficiency levels, but it is less clear whether these factors (in particular strategy use) constitute learner IDs or reflect the listening proficiency level as such. Nevertheless, experimental studies have demonstrated that IDs such as metacognition for L2 listening can be enhanced through intervention. Vandergrift and Tafaghodtari’s (2010) process-based approach—involving predictions, verifications, and reflections on listening text content and comprehension—is one example of a successful instructional treatment that could be adopted in teaching and learning contexts. Future research should, however, continue to study the relationship between ID factors and L2 listening success, so that language teachers, language testers, and policy makers can attune their practices to the benefit of the learners.

**Future Directions**

A consistent observation across all IDs reviewed in this chapter is the scarcity of empirical research specifically focusing on ID factors in L2 listening. This applies to the investigation of IDs themselves (as relevant to listening), and even more so to their relationship with and impact on L2 listening. There is therefore a large scope for research in this area, across the board, but perhaps mostly so in terms of conative difference factors in L2 listening since hardly any research (or sometimes none at all) was found on motivation, mindsets and regulatory focus (see Chapter 8), goal complexes (see Chapter 9), or willingness to listen in relation to L2 listening development and performance. There
also appear to be no or very few studies specifically focusing on L2 listening and the cognitive differences of language aptitude (see Chapter 2) and declarative/procedural memory (see Chapter 4), and on L2 listening and the affective difference of enjoyment (see Chapter 12). Researchers should focus on these areas to further our understanding of what contributes to successful listening. Ideally, however, this should not be done through studying factors in isolation, but by employing research designs aimed at investigating how they interact (see also the discussion below). A related recommendation is that, as the body of research builds up, it will be useful to start conducting replication studies to evaluate the generalizability of findings on ID factors in L2 listening across contexts, languages, and populations.

A second observation is that much of the existing literature on IDs in L2 listening does not seem to be grounded in a clear theory of the ID factor under scrutiny, nor in a clear theory of (L2) listening itself. This may be due to the newness of studies in the area; however, it is vital that scholars clarify what theoretical models frame their work in studying how and why individuals differ with regard to the rates, routes, and levels of ultimate attainment in L2 listening. Without substantiated and explicit construct definitions, and representative operationalizations, research findings on IDs in L2 listening are not interpretable, not comparable across studies, and not implementable in the teaching, learning, and assessment of L2 listening.

A third important conclusion from the above review is that relatively little is currently known about the interaction between different ID factors (e.g., metacognition and self-regulated learning, or motivation and anxiety; but see e.g., Bang & Hiver, 2016; Nix, 2021) and their relative contributions to L2 listening. Similarly, few studies have investigated the interaction between ID factors and other well-established factors such as L2 vocabulary knowledge (or other task and listener characteristics), and their relative contributions to L2 listening (but see e.g., Andringa et al., 2012; Bang & Hiver, 2016). To increase our insights in this respect, future empirical studies will need to go beyond “simple” correlational analyses and adopt more complex analytic methods such as structural equation modeling (see Bang & Hiver, 2016; Chapter 32), factor analyses (see Chapter 32; Shepherd & Ainsworth, 2017, 2018a, 2018b), or mixed-effects modeling (for an example from the area of L2 writing, see e.g., Michel et al., 2019). Such sophisticated statistical analyses will also require higher participant numbers than have been involved in many existing studies. For example, a future study could investigate the interaction between declarative/procedural memory, self-efficacy, and enjoyment in determining listening success on a large sample (N > 200) of L2 learners, employing mixed-effects modeling. Such a study would have the potential to reveal important relationships between different ID factors. For instance, it may be found that listeners’ declarative/procedural memory influences their self-efficacy, which in turn may also lead to differences in enjoyment of L2 listening (e.g., higher declarative/procedural memory leads to higher self-efficacy, which results in greater enjoyment). Such findings would help practitioners make more informed decisions about their pedagogical choices in teaching and assessing L2 listening.

Another recommendation for researchers intending to explore ID factors in L2 listening concerns the need to validate their research instruments with respect to their intended use and score interpretations, and with respect to the intended research population. While an existing measure and the interpretations of the scores on that instrument might have shown to be valid in an earlier published study, this does not automatically carry forward to other research contexts, purposes, and populations. It is therefore vital that researchers—whether they develop their own instruments, or select or adapt an existing instrument—gather evidence that supports the accuracy of score interpretations in their particular setting through rigorous reliability and validity testing. After all, if there are validity concerns associated with the instruments and their uses in a particular study, the conclusions drawn from that study about ID factors in L2 listening will lack credibility. For example, when using an existing L2 motivation questionnaire in a new study, it is crucial to test the underlying assumptions of the questionnaire using statistical methods, such as factor analysis and
reliability testing. An introductory guide on validating L2 instruments, their uses, and interpretations for research purposes can be found in Révész and Brunfaut (2021).

Finally, as the research base on ID factors in L2 listening becomes more comprehensive, elaborate, grounded, and methodologically sound, it will be meaningful to similarly expand research into practical applications for the development of learners’ L2 listening, informed by the empirical insights into ID factors and L2 listening. This could comprise experimental studies (e.g., interventions with pre-testing and post-testing) on pedagogic techniques targeting IDs that have been shown to affect or interact with the teaching, learning, or assessment of L2 listening.

References


Franz Holzknecht and Tineke Brunfaut
IDs for L2 Listening


