This chapter examines four individual differences (IDs) in relation to learners’ L2 speaking, which I broadly define as spontaneous production in oral face-to-face (FTF; Trebits, 2016) or oral synchronous computer-mediated communication (CMC; Satar & Özden, 2008) contexts. The research investigating links between learner IDs and second language (L2) speech has often focused on L2 speaking as a product, overall, more so than with respect to L2 speech development or as a response to feedback; the emphasis in the current chapter follows suit. The four IDs, one from each of the sub-categories of IDs within this volume, were selected given their prominence as the most-researched ID within each category with respect to learners’ L2 speaking.

The Role of IDs in SLA

First, it is important to mention that theories of second language acquisition (SLA) posit a differentially important role for learners’ L2 speaking. Some, such as the cognitive-interactionist approach (Long, 1996), maintain that learners’ L2 production is critical for learning to occur, as it offers learners the chance to test hypotheses, receive valuable and necessary feedback with respect to the target-like quality of their L2 production and, as a result, trigger adjustments to their interlanguage and/or output. Sociocultural approaches (Lantolf, 2020) describe the centrality of L2 production as allowing learners to approximate and adapt the expertise of surrounding interlocutors, their zone of proximal development; importantly this can include physically present interlocutors as well as past, future, or even imagined communicative partners. The importance of spontaneous L2 production, which is arguably the majority of L2 speaking, is viewed to be important in several theories as it pushes learners to stretch beyond their comfort zones and automatize their skills (output hypothesis, e.g., Swain, 2000; skill acquisition theory, e.g., DeKeyser, 2015). Theories also differ with respect to their view on how learners’ IDs play a role in their L2 development. For an extensive cross-theoretical discussion of the roles of learner IDs, the reader is directed to Chapter 1 (see also Gurzynski-Weiss, 2020).

Given these theoretical variances, we would expect a differential empirical weight in investigations of learners’ IDs with respect to their L2 speaking abilities and this is indeed what we find. In the sections that follow I focus on the ID from each of the categories of this volume that has been the most researched with respect to learners’ L2 speaking, allowing for optimal demonstration of how learner IDs along each dimension influence L2 speaking and for cross-chapter discussion and
tie-in. First, I will briefly reiterate the theoretical importance of each ID; in-depth discussions are discussed elsewhere in this volume.

**Age**

The age of onset, or the age at which learners begin to learn an L2, is important for all theories of SLA. In a most basic distinction, learners who begin to acquire their L2 before puberty, a “critical period” when the brain specializes (Lennenberg, 1967), will learn the L2 to a higher level than counterparts who start after puberty. While details are beyond the scope of the current chapter (see Chapter 16), there is evidence that there are multiple critical periods, with the first for accent closing at around age six (Walsh & Diller 1981), and grammar and lexis following later. Age matters not just for physical language-devoted brain space or processing abilities; when learners begin to acquire an L2 earlier there is a greater likelihood that they are acquiring the L2 outside of the classroom, where input is rich, authentic, and varied, as are opportunities to interact with a wide range of interlocutors. Thus, learner age often overlaps with learning context, L2 opportunities, as well as learner cognitive abilities.

**Working Memory**

While there are numerous definitions, working memory (WM) can be defined as the memory at play when learners are carrying out tasks, including L2 speaking, which requires simultaneous processing, retention, retrieval, and use of information. As discussed in Chapter 3, there are two perspectives on the role WM plays in L2 speaking. The first, the cognitive-interaction approach, maintains that WM enables learners to notice and process input received during interaction (Mackey & Gass, 2015). Similarly, Levelt’s (1992) speech model views WM as critical at the monitoring stage, when learners attend to what others are communicating, as well as the conceptualization stage, where learners’ attention and memory resources are activated when deciding upon message creation. Thus, WM makes L2 speaking possible; it allows learners to process the input that they receive and to parse and momentarily store the information while simultaneously formulating a response.

**Anxiety**

Anxiety is both “an emotional reaction and disruptive to ongoing cognition and behavior—thus making it part of a continuous cycle of sometimes influencing other variables and sometimes being influenced” (Gregersen, 2020, p. 79). Anxiety types include trait (e.g., I am an anxious person in general; Spielberger, 1983), state (e.g., I am anxious right now completing this task; Baralt & Gurzynski-Weiss, 2011), or situational, which occurs whenever someone is in a certain situation (e.g., I am anxious when I speak in public; Ellis, 1994). There is also foreign language classroom anxiety, which is situationally specific to using the L2 in the classroom (Horwitz et al., 1986). Critically for L2 speaking, anxiety is extremely dynamic, changing in moment-to-moment interactions with the potential to significantly reduce or eliminate learners’ attention to the processing and/or production of spoken L2. When anxiety reaches a certain level (for each individual), it can interfere with one’s ability to produce the L2 and, at certain levels (idiosyncratic; see Chapter 11), negatively impact the ability to process the L2 received.

**Willingness to Communicate**

Willingness to communicate (WTC) is the probability that a learner chooses to speak when given the opportunity (MacIntyre, 2020). Thus, for the interaction approach, sociocultural theory, and
complex dynamic systems theory, WTC is of central importance, as learners’ interaction with others, including L2 production, is critical for SLA. WTC is not a concern for universal grammar or the input hypothesis, which view learners’ reception of the L2 as central.

Research

Evidence

In this section, I focus on evidence of the four IDs in relation to learners’ L2 speaking production and, at times, development. I provide both a narrative synthesis of trends and specific examples to illustrate the nature of the existing work.

Age

Age is a sociocultural/demographic ID (Chapter 16). Historically, age has been defined as learners’ number of years post-birth, or their age at the onset of learning a particular language, which is binarily categorized as pre- or post-puberty. More recently, work by Pfenninger and Singleton (2017) has recognized how age effects—most notably the effects of age of onset—are mediated by contextual and additional IDs. Thus, while some operationalizations (i.e., number of years post-birth) are dynamic, changing with the passage of time, age can also be considered as dynamic in terms of its influence given additional variables at play. Other operationalizations of age, such as the age of onset of learning a particular L2, are static.

Despite substantial research on age as an influencing factor in SLA, there is little research focusing on potential relationships between learners’ age and their L2 speaking beyond accent. In this section, I detail representative studies investigating relationships between age and learners’ negotiation of meaning and feedback, oral production on specific tasks, and examinations of L2 speaking abilities in relation to learners’ age.

Learners’ Age and Negotiation of Meaning

First, differences in negotiation of meaning (NoM), or a mutual attempt to arrive at understanding between two or more people, have been found to relate to learners’ age. Operationalizing age as the years post-birth, García-Mayo and Lázaro Ibarrola (2015) examined potential differences in NoM and L1 use for 40 8–11 year-olds enrolled in either content and language integrated learning or English L2 classes. Using a spot-the-difference picture task, they found older learners, regardless of instructional context, negotiated less and used the L1 more frequently than younger learners. Using a jigsaw task, Hidalgo (2019) observed differences in negotiation strategies used during learners’ oral task completion. Forty 8–9 year-olds and 40 10–11 year-olds had the same range of NoM strategies. However, the younger learners used clarification requests and self-repair, whereas the older learners acknowledged understanding (e.g., ok, yes).

Also, operationalizing age as the age at the time of study, Oliver et al. (2017) examined multiple facets of L2 speaking, including social interaction (cooperation, reciprocity, and resolution of conflict), task management (organizing the task and on-task talk), cognitive involvement (cognitive talk and repetitive patterns), and focus on language (supporting, language production, or play). Forty-two children from four L2 English classes participated; 11 pairs of children ages 5–7 and ten pairs aged 11–12. Oliver et al. discovered older learners showed more cooperation and reciprocity and engaged in repetitive patterns and task management, while younger learners demonstrated conflict resolution and engaged in language play. Thus, even in three examples, we see learners’ engagement, production, and response to L2 speaking differs modestly according to the age when operationalized as years old at the time of study.
Learners’ Age and L2 Accent on Specific Tasks

Most studies examining potential relationships between learners’ age and L2 pronunciation have used specific spoken production tasks rated by experts in L2 phonology or native speakers (NSs). For example, operationalizing age as learners’ first intensive exposure to an L2 environment (age of acquisition [AoA]), Saito (2015) examined the accentedness (nativeness) and comprehensibility (ease of understanding) of 88 L2 English learners (L1 Japanese) living in Canada while completing a timed description of three pictures, judged by ten English NSs and using the data of ten NSs of L1 Japanese and L1 English as baselines. He also completed additional segmental, prosodic, temporal, lexical, and grammatical analyses. Saito found a significant relationship between AoA and accentedness and comprehensibility when the length of residence was controlled for using partial correlations: the greater AoA, the greater accentedness and the lesser comprehensibility. Saito also reported significant positive correlations between AoA and word stress and intonation; AoA was not correlated with vocabulary and grammar. Thus, Saito confirms previous research that phonology is more sensitive to age effects than vocabulary or grammar (Granena & Long, 2013; DeKeyser & Larson-Hall, 2005) and, critically, that AoA is influential for later SLA.

Looking at the relationship between the age of onset (AO; a synonym of AoA) and voice onset time (VOT) within the realm of accentedness, Stölten et al. (2015) examined the speech of 41 L2 Swedish (L1 Spanish) near-native speakers during a word-reading task. AO was operationalized binarily as either early (before age 12) or late (after age 12); 31 were early AO learners and ten were late. Specifically, they investigated learners’ production of /ptk/, given that Swedish has longer VOTs than Spanish. The researchers discovered that 71% of the early AO learners produced more native-like Swedish VOT in all three stops, 16% produced native-like VOT for two stops, and 10% produced native-like VOT for one stop, whereas only 50% of later AO learners produced native-like VOT for all three stops, 20% produced native-like VOT for two stops, and 30% produced native-like VOT for one stop. In other words, like Saito (2015), Stölten et al. (2015) found AO to directly relate to accent in L2 speaking: greater AO relates to greater accent, at least for Swedish /ptk//VOT.

In a study of 65 L2 Spanish (L1 Chinese) learners living in Spain (with 12 NSs of Spanish as a baseline), Granena and Long (2013) similarly reported a significant correlation between AO and pronunciation, operationalized as the degree of foreign accent while reading a three-line paragraph, rated by 12 linguistically naïve NSs of Spanish. Dividing the learners into three groups, AO 3–6 years, 7–15 years, and 16–29 years, Granena and Long found only the youngest AO learners (only those aged five and younger) to score in the native range. Age and accent correlated in the 7–15 group and flattened out in the 16+ group. This study adds to the prior two in supporting age effects (in terms of the onset of SLA) for pronunciation during L2 speaking, providing evidence that the optimal window is five or younger, and that after 15 one’s opportunity plateaus.

Focusing on the age of instruction and age of immersion, Moyer (1999) examined a potential relationship between the nativeness of speech and age for 24 L2 German speakers (L1 English). Participants read a list of 24 words, eight sentences, a paragraph, and a spontaneous speech sample, and their speech was judged by four NSs of German. Moyer also measured learners’ motivation for taking German, perceived L2 pronunciation abilities, beliefs regarding the importance of native-like pronunciation, and asked about the type of L2 instruction they received (amount, immersion, and phonological instruction). The age of instruction correlated negatively with pronunciation ratings: the later the instruction, the less native-like the perceived accent. A negative correlation was also observed for immersion: the higher the age of first immersion, the lower the native-like rating the learner received. Moyer thus provides additional evidence that earlier is better for L2 speech.

In a final example of operationalizing age as the age at the time of recording (controlling for no exposure prior to age 12), Bongaerts et al. (1997) examined a potential relationship between age and ultimate attainment, operationalized as ratings of nativeness (1 = definitely nonnative,
IDs for L2 Speaking

5 = definitely native), which were examined in two studies. In the first, ratings of nativeness were given by four inexperienced raters (no linguistics training) who heard learners produce spontaneous speech samples and read a text of 84 words, ten sentences, and 25 English words. No statistical differences were uncovered between the NS baseline of British English and the ten L2 English learners (L1 Dutch) who received intensive instruction on pronunciation and spent at least a year in Britain. In the second study, 13 judges, six of whom were experts (EFL teachers/phoneticians) and seven of whom were inexperienced, heard learners read six sentences. Five of the 11 L2 English speakers (L1 Dutch) who spent at least a year abroad and had pronunciation instruction were rated as having native-like pronunciation, although on only four of six sentences; for all other measures, learners’ greater age directly related to an increase in perceived accent.

Thus, for L2 pronunciation, at least when operationalized as accent, there is a relationship with age: the later learned, the more accented the speech. There is also evidence (Bongaerts et al., 1997), however, that spending a year abroad with instruction can help. Additionally, there are differences in how learners negotiate meaning based on their age. This area of research is particularly ripe for additional work, as L2 speech is much greater than simply perceived accent, and age, as argued by Singleton and Pfenninger, is not an isolated variable.

Working Memory

The cognitive ID that has been most investigated with respect to L2 speaking is WM. WM is a dynamic ID in the sense that it is believed to change over one’s lifetime (see review in Jackson, 2020), and its role and influence changes as learners’ proficiency and/or age increases (for empirical studies see Sanz & Serafini, 2018; for a theoretical overview see Wen, 2015). Specifically, WM is stronger in adults than children until it begins to decrease again in the later years. However, compared with other IDs, WM is less dynamic than WTC and anxiety, which can change moment to moment, and has been typically investigated at one point in time with respect to L2 speaking. In their (2014) meta-analysis examining WM and L2 comprehension and production, Linck et al. reported WM to positively correlate with L2 processing and proficiency gains, echoing Watanabe and Bergeleithner’s (2006) meta-analysis. In terms of individual studies investigating the link between WM and L2 speaking, we find examples investigating relationships between WM and learners’ self-repair and modified output, between learners’ WM and learners’ speaking as conceptualized by specific measures (most often complexity, accuracy, and/or fluency), and relating WM to gains and abilities in L2 speaking and/or measures of general L2 speaking proficiency. Each of these branches is discussed in turn.

Working Memory and Learners’ Self-Repair and/or Modified Output

As predicted within the monitoring portion of Levelt (1992), WM has been documented to relate to learners’ self-repair (the repair they make when hearing their own mistakes) as well as their modified output (modifications made to their spoken L2 production following feedback, verbal or non-verbal, from communicative partners). For example, Mackey et al. (2010) examined verbal WM in relation to the modified output of 42 undergraduates enrolled in fourth-semester Spanish produced during four oral tasks (map, picture-drawing, spot-the-difference, and story completion tasks). Learners with higher levels of (composite) WM produced more modifications than learners with lower WM (see Chapter 3 for WM measurement discussion). When recall and processing components of WM were separated out, only the recall component was significantly related to modified output. A positive relationship between WM and learners’ production of modified output was also found in Bergeleithner (2011), who also utilized 42 undergraduates of fourth-semester Spanish. Like Mackey et al., Bergeleithner found learners with higher WM produced more modified output. Focusing on learners’ self-repair, and operationalizing WM via a listening
span task, Mojavezi & Ahmadian (2014) had 40 intermediate-level Iranian L2 English workers complete two oral narrative tasks, one in their L2 English and one in their L1, Farsi; they then conducted a stimulated recall to better understand learners’ self-repairs. Positive correlations were found between WM and L2 repair behavior and not found between WM and L1 repair, lending support to the mediating factor of proficiency level mentioned earlier (see also Wen, 2015). Thus, in these studies, WM mediated learners’ ability to modify their output, which demonstrated a cognitive registration of a gap between their original and target utterances, and visible effort to close the gap.

**Working Memory and Learners’ L2 Spoken Performance: Complexity, Accuracy, and Fluency**

Positive relationships have been uncovered between WM and L2 spoken production operationalized as complexity, accuracy, and/or fluency (CAF) at one moment in time. Looking at all three CAF measures, Borges Mota (2003) had 13 advanced learners of L2 English complete picture description and narrative tasks. Their speech was analyzed in terms of complexity (number of dependent clauses per minute), accuracy (number of errors per 100 words, counted across syntax, morphology, and lexical choice), and fluency (speech rate—pruned and unpruned—as well as number of silent pauses and hesitations per minute, mean length of run). He also examined learners’ L2 speech in terms of weighted lexical density; high-frequency items appearing more than once in the sample were given half the points as low-frequency items. Borges Mota reported positive correlations between WM and fluency (speech rate and mean length of run fluency), WM and complexity, and WM and accuracy, and unexpectedly negative correlations between WM and weighted lexical density for the picture description task; for the narrative task no relationship was found between WM and lexical density, demonstrating a mediating effect for task type. Borges Mota suggests these results lend support to Skehan’s trade-off hypothesis, demonstrating that when learners increase their CAF, they are unable to provide lexically dense speech.

Ahmadian (2012) reported a similar relationship between 40 intermediate-level male L2 English Iranian learners’ WM, measured by a listening span task, and L2 performance on an oral narrative of a 14-minute silent cartoon. Positive relationships were found between learners’ L2 speaking accuracy (error-free clauses and correct verb forms) and fluency (number of syllables/minute and number of meaningful syllables per minute after removing repetitions, reformulations, and replaced words). No relationship was uncovered between their WM and L2 speech analyzed for complexity, operationalized as both syntactic complexity (amount of subordination) and syntactic variety (different grammatical verb forms used). Weissheimer and Borges Mota (2009) reported significant correlations between measurements of accuracy (number of errors per 100 words and percentage of error-free clauses) and complexity (number of dependent clauses per minute, number of subordinate clauses per T-unit, and number of clauses per C-unit) produced via L2 speech in a picture description task completed by 32 L2 English learners; WM was measured by a speaking span test. No relationship between WM and fluency (number of words per minute) was observed. Kormos and Trebits (2011) found task complexity to mediate the relationship between WM (backwards digit span test, involving the central executive and phonological loop) and learners’ accuracy (error-free clauses and verbs), speech rate, and clause length for 44 secondary school learners of L2 English (L1 Hungarian). Specifically, a relationship between WM and learners’ accuracy was uncovered only for the less-complex oral picture narrative task where the pictures related to a common storyline. No relationship between WM and accuracy was found for the more complex narrative task where learners had to invent a story from unrelated pictures, and no relationship was found between WM and fluency (speech rate and number of syllables produced divided by time spent speaking), grammatical complexity (average length of clauses and number of different verb forms), or lexical complexity. In a final example, Finardi and Prebianca (2006) discovered a signifi-


IDs for L2 Speaking

A significant relationship between a WM speaking span task and 12 intermediate L2 English (L1 Brazilian Portuguese) learners’ fluency (unpruned speech rate) when they completed a picture description task. No relationships were found between WM (operation-word span task), and neither measure of WM was significantly correlated with complexity (number of dependent clauses per minute), accuracy (number of errors per minute), or lexical density (weighted rate of use including low-frequency and high-frequency words). Thus, WM has been consistently discovered to be related to L2 speaking for one or more CAF measurements, with higher WM beneficial for one or more L2 speaking components.

This finding, that WM relates to one or more of the CAF measures but inconsistently so, is echoed in studies examining a single aspect. For example, focusing on accuracy in L2 speech, Fortkamp and Bergsleithner (2007) used a WM speaking span task and had 18 pre-intermediate L2 English learners of L1 Brazilian Portuguese complete an oral picture description task. They observed a positive relationship between WM and accuracy when operationalized as the number of errors per 100 words. An earlier study by Fortkamp (1999) featured 16 L2 English learners of L1 Portuguese and examined fluency, measured by oral performance on the picture description and reading aloud tasks, and an oral slip task eliciting spoonerisms (“the transposition of phonemes in adjacent or near adjacent syllables or words”; Motley, Baars, & Camden, 1983, p. 81). Fortkamp examined these measures of fluency in relation to learners’ WM (a speaking span test administered in learners’ L1 Portuguese and L2 English, as well as a reading span test given in both languages). Fortkamp reported that the English L2 speaking span test correlated significantly with the picture description, and that both the L1 and L2 reading span tasks correlated with time-on-task during the oral reading (i.e., the greater the WM, the faster the learners read aloud). Again, we see a positive relationship between WM and L2 speaking when only examining one specific aspect of L2 speech, in this case reading aloud.

Working Memory and Learners’ L2 Speaking Proficiency

WM can also relate to L2 speaking when measured holistically at one point in time. For example, Finardi and Weissheimer (2008) observed learners with a higher level of L2 speaking proficiency (based on their description of a picture-cued narration, rated by experienced raters) to have greater WM (L2 speaking span test). The learners in their study were 79 L2 English learners in Brazil, 47 who were lower proficiency and 32 who were higher. The authors cautioned conclusions, stating it was impossible to say whether the learners’ WM helped them with their comparatively greater proficiency (e.g., causation), or if learners with higher WM were simply more likely to continue on with their L2 studies, so the sample represented those who chose to continue on in their language study.

Positive correlations have also been found between learners’ WM and L2 speaking development measured longitudinally. For example, O’Brien et al. (2006) discovered phonological WM (tested via serial order recognition of lists of nonwords) to relate to productive vocabulary (unique words), narrative abilities (past tense, third person morphology, past participle, present participle, and communication verbs), free grammatical morphemes, and subordinate clause use at both Time 1 and Time 2 (13 weeks apart). O’Brien et al. (2007) conducted additional analyses from this sample, examining four-minute extracts from the earlier Oral Proficiency Interview (OPI) and calculated learners’ general oral ability (total number of words produced and the number of words in the longest turn) and measures of fluidity (speech rates in words/minute, absence of hesitations, absence of filled pauses, and longest fluent run). O’Brien et al. reported phonological WM to be significantly related to the absence of filled pauses and the absence of hesitations at Time 2. In terms of development, phonological WM was significantly related to residualized change scores in learners’ longest fluent run. In other words, the greater the (phonological) WM, the more fluent the speech and the greater the change in oral ability.
Finally, an inconsistent link has been discovered between WM and proficiency gains following practice via CMC, a mode which has been theoretically posited to benefit learners who have lesser WM, as there is greater visual support and learners may need to rely less on WM to successfully engage. In their 2002 study examining if CMC builds oral proficiency and if gains are mediated by WM, Payne and Whitney used a recognition-based nonword repetition task and reason span measure for WM and operationalized proficiency as rated performance (with two raters) on a five-minute speech sample. Utilizing 58 learners of L2 Spanish enrolled in third-semester courses, they found a significant relationship between the nonword repetition task and proficiency gain scores. In a follow-up study, Payne and Ross (2005) used a recognition-based nonword repetition task (a measure of phonological WM) and reading span measure (a measure of executive function) with a five-minute oral speech sample as the measurement of L2 speaking. This time with 24 learners of Spanish in a third-semester course, they found a positive relationship between phonological WM (PWM) and oral proficiency gains over a semester; meaning that higher PWM learners had more oral gains than lower PWM colleagues; this positive relationship was not found for the reading span. Thus, it appears that WM mediates processing and output differentially in the CMC mode than in person.

In summary, the relationship between WM and learners’ L2 speaking has been observed to be largely positive: overall, the greater the WM, the more complex (e.g., Borges Mota, 2003), accurate (e.g., Weissheimer & Borges Mota, 2009), or fluent (e.g., O’Brien et al., 2007) learners’ speech is. This relationship appears to be mediated, however, by L2 proficiency (e.g., Finardi & Weissheimer, 2008), the mode of communicative practice (e.g., Payne & Ross, 2005), complexity (Kormos & Trebits, 2011), and task type (e.g., Borges Mota, 2003).

Anxiety

As detailed in Chapter 11, anxiety is an affective ID with the potential to influence each moment of learners’ engagement with L2 speech—input, processing, and output (Bailey et al., 2000). It is inarguably one of the most researched affective IDs in L2 learning, as described in the 2019 meta-analysis by Teimouri et al. (see also Gkonou et al., 2017; Gregersen, 2020). In contrast to the positive relationship between WM and speaking, the relationship between learners’ anxiety and L2 speaking is overwhelmingly negative: the higher learners’ anxiety, the less likely they are to produce target-like L2 speech. In the paragraphs that follow, I detail studies that have examined potential links between learners’ anxiety and their L2 speech production on specific tasks.

Anxiety and Learners’ Processing of Spoken L2 Input and Production of Spoken Output

Focusing on how learners decode and produce L2 spoken narration, Trebits (2016) used MacIntyre and Gardner’s (1994) L2 input, processing, and output anxiety questionnaire along with cartoon oral descriptions and picture narratives analyzed according to complexity (clause length and ratio of subordinate and relative clauses), accuracy (ratio of error-free clauses, relative clauses, and error-free past-tense verbs), and fluency (speech rate). Trebits, focusing on 44 L2 English learners (L1 Hungarian) enrolled in the second year of a bilingual program, discovered anxiety levels to be rated higher on the questionnaire for processing spoken input than for processing input presented in aural or written modes. While no relationships were found between input anxiety and oral production, processing anxiety correlated with complexity (lexical diversity and ratio of subordinate clauses) for the oral cartoon description, and output anxiety correlated negatively with accuracy (error-free clauses and past-tense verbs) in the oral story narration. Thus, like studies reviewed earlier for WM (Borges Mota, 2003), task type mediated the relationship between anxiety and L2 speaking.
Anxiety and Learners’ L2 Speech Production

Learners’ anxiety has been reported to relate negatively to L2 oral production when completing a variety of tasks. First, we will review research investigating potential relationships between anxiety and L2 speaking that has operationalized anxiety as L2 classroom anxiety, often employing the foreign language classroom anxiety scale (FLCAS) questionnaire. For example, in an earlier study, Phillips (1992) found a negative correlation between the FLCAS scores of 44 learners of third-semester L2 French and their graded performance on a two-part oral exam—a monologue on a randomly selected topic—and their performance in three role plays. In their 2008 study, Satar and Özdener had 90 L2 English learners (L1 Turkish) complete a 20–30 minute dialogic speaking task before and after a four-week intervention (two tasks per week completed in 40–45-minute-long dyadic chat sessions: info-gap, problem-solving, jigsaw, and decision-making) with either voice or chat-based CMC practice; one-third of the group served as a control group, completing the speaking task four weeks apart without intervention. Learners’ oral L2 speaking was coded for general proficiency (pronunciation, vocabulary, grammar, fluency, and understanding) and course-specific measures (content and skills, vocabulary, and grammar) as well as the FLCAS. The voice chat group had a larger increase in speaking proficiency, though their anxiety did not change significantly. Correlations between anxiety and L2 speaking showed that anxiety and speaking scores were negatively related at the beginning of the study, but not after four weeks. Thus, for this study, learners’ L2 classroom anxiety was effectively managed by the intervention, so much so that it no longer impacted their L2 speaking.

Let us turn to two example studies investigating non-L2 anxiety in relation to specific components of L2 speaking. In Woodrow (2006), 275 L2 English learners’ reported in-class and out-of-class anxiety in relation to a ten-minute speaking task conducted in three stages: an introduction and general interview, individual statement, and a two-way discussion. Learners’ L2 production was assessed for fluency, language usage, and pronunciation. Woodrow found a positive correlation between in-class and out-of-class anxiety, and a negative relationship between oral performance and in-class anxiety. In a second study examining non-L2 anxiety and aspects of L2 speaking, Oya et al. (2014) utilized the Spielberger state-trait anxiety inventory to examine 73 learners of L2 English (L1 Japanese) in New Zealand language schools and found anxiety to negatively correlate with learners’ accuracy (for clauses and verbs) in their description of six picture cards from the Wechsler (1966) adult intelligence scale. No relationship was observed between anxiety and learners’ complexity (number of words per T-unit and mean length of three longest utterances), fluency (number of syllables per second excluding fillers and discourse markers like “I don’t know”), nor a global impression rated by three individuals. For Woodrow (2006), much like the studies reviewed above (Phillips, 1992; Satar & Özdener, 2006), L2 situational anxiety focusing on “being in a classroom” negatively correlated with oral performance; a negative relationship was also found in Oya et al. (2014) for fluency.

Anxiety and Learners’ Self-Rated L2 Proficiency

Finally, we review two studies that have examined the associations of anxiety with learners’ self-rated proficiency measures. Using the 1988 version from MacIntyre and Gardner’s L2 input, processing, and output anxiety questionnaire, MacIntyre et al. (1997) reported 37 L2 French (L1 English) Canadian learners’ anxiety negatively correlated with the overall quality of their output (a composite of fluency, complexity, accented speech, elaboration, grammar, and similarity to Francophone colloquial expressions), and self-rated proficiency (rated via can-do statements). Piechurska-Kuciel (2011) also found a negative relationship between 393 L2 English learners’ FLCAS scores and self-assessed L2 speaking ability (one item, rated on a six-point Likert scale from very poor to excellent), collected once per year over three years. Thus, anxiety, whether measured via the input, processing, and output anxiety questionnaire (MacIntyre & Gardner,
negatively correlates with learners’ self-proficiency, measured at one point in time (MacIntyre et al., 1997) or over three years (Piechurska-Kuciel, 2011).

In summary, throughout this section it is clear that learners’ anxiety, whether it be L2-specific or anxiety beyond the L2, negatively relates to their L2 speaking abilities in terms of complexity, accuracy, and fluency (MacIntyre et al., 1997; Trebits, 2016); more holistic proficiency measures taken at one time (Woodrow, 2006); measuring gains (measuring L2 speaking at more than one time point; Piechurska-Kuciel, 2011); and self-rated proficiency (MacIntyre et al., 1997). Let’s now turn our attention to the final ID in this chapter, WTC.

**Willingness to Communicate**

As detailed in Chapter 10, WTC is a conative ID, meaning that it deals with the domain of volition and motivation. While originally conceptualized as a more stable personality trait, WTC has since evolved to be understood to fluctuate greatly in moment-to-moment interactions (see MacIntyre, 2020). Additionally, L2 WTC is conceptualized differently from L1 WTC, though they are often related; those with higher L2 WTC are more likely to also have higher L1 WTC, though the reverse does not necessarily hold. Thus, much like anxiety, while a given L2 learner can be considered to have comparatively greater or lesser L2 WTC than others, WTC is dynamic and changing based on the myriad factors at play in each individual interaction where the opportunity for L2 speaking exists.

**WTC and Learners’ L2 Speaking in the Classroom**

Several studies have examined WTC in relation to spoken production in L2 classrooms. For example, in Munezane (2016), 373 students of L2 English in Japan completed a 20-item instrument adapted from Sick & Nagasaka (2000), rating how often they would communicate in different situations on a four-point Likert scale. This measure of WTC was examined in relation to the number of words each student produced in three classroom group discussions. A significant relationship was documented between learners’ L2 use during the three class discussions and their L2 WTC, meaning that learners who reported higher WTC spoke more than those with lower WTC.

Bernales (2016) also examined L2 WTC in relation to L2 classroom participation, although in this study, the 16 learners of third-semester L2 German (L1 English) rated their own participation. Using a questionnaire developed specifically to investigate “WTC as it originated in students’ minds during class” (p. 4), students were asked to predict how much of their participation would be in German; after 25 minutes students were asked to respond with how they actually participated (repeated for four class meetings). Four focal students completed a 50-minute interview and had their WTC actions measured via classroom observations. Bernales reported students’ predictions to become more aligned with their actual WTC. In addition, learners’ L2 use increased as they reported they were more engaged in the course as time went on. Thus, both studies (Munezane, 2016 and Bernales, 2016) found learners’ WTC to relate to participation in their L2 classroom, whether rated by a researcher (Munezane, 2016) or by learners (Bernales, 2016). Bernales (2016) additionally found that learners’ participation could be increased by making them aware of and charting their WTC and participation during a classroom lesson.

**WTC and L2 Speaking During Specific Tasks: The Idiodynamic Method**

Two studies using the idiodynamic method (MacIntyre, 2012) have found relationships between WTC and L2 speaking during specific tasks. The idiodynamic method allows the researcher (and/or teacher) to elicit multiple perspectives of a single interaction (see Chapter 32). Relying on a videotaped interaction, participants watch the video and rate (prompted and/or unprompted)
their or another’s (perceived) WTC in each moment. This method allows nuanced examination of a single interaction and permits the discovery and measurement of WTC on a moment-to-moment basis. For example, in a case study of four L2 English learners (L1 Japanese) in Canada, Wood (2016) examined fluctuating WTC in relation to speech rate, pause-time ratio, and length of fluent runs between pauses during a picture description task. Discussing the results in terms of “waves” of influence between higher/lower WTC and higher/lower measures of the aforementioned aspects of L2 speaking, Wood found five waves of low WTC leading to low fluency, four waves of low fluency leading to low WTC, four waves of high WTC leading to high fluency, and one wave of high fluency leading to high WTC. Thus, the idiodynamic method uncovered the interrelatedness and causality between WTC and fluency which would have been obscured by less-dynamic methods. In another study using this methodology, Nematizadeh and Wood (2019) found a relationship between the WTC of four L2 learners of English (L1 Farsi) and their mean length of run. Specifically, the researchers reported that, in general, learners’ higher WTC translated to a higher mean length of runs in their L2 speaking during a picture description task, and found lower WTC to be associated with shorter, fragmented segments and frequent pauses. When WTC and mean length of runs interacted negatively, the researchers reported learners used prefabricated chunks or, alternatively, had too many ideas they wanted to express, which inhibited their decision to act in that moment. The takeaway from both studies is that the idiodynamic method permits the discovery of the relationship between WTC and L2 speaking as well as revealing the why behind it.

WTC and Mediating Variables

Finally, several studies have examined potential mediating variables within the relationship between WTC and L2 speech. Exploring a potential link between L2 fluency and learners’ WTC and considering the mediating variable of the context of instruction (at-home as compared to a short-term study abroad), D’Amico (2012) had 23 learners complete self-rated questionnaires of WTC (based on MacIntyre et al., 2001) for their L2 Spanish and L1 English; nine learners participated in the study abroad (SA) and 14 were studying at home (AH). Whereas D’Amico found a relationship between SA participation and increased speech rate, average length of fluent runs, and self-repairs, AH learners showed a decrease in unfilled pauses and dysfluent clusters. No relationship was observed for either group between WTC and L2 fluency measures, which were based on two one-minute extracts from pre- and post-program oral interviews between each participant and the author. In another study examining a potential mediating factor between WTC and L2 speech, Yanguas and Flores (2014), investigated if the mode of interaction (oral CMC or oral FTF) was a mediating variable for WTC (measured with a questionnaire from Cao & Philp, 2006) for 31 intermediate-level learners of L2 Spanish (L2 English) completing an oral discussion activity in their L2 classes. They reported a significant relationship between WTC and the number of turns and words in FTF only.

Thus, regardless of the particular L2 speaking measure or task type, overall, there seems to be a positive relationship between learners’ L2 WTC and L2 speech. This is especially evident in the FTF mode and for WTC that is conceptualized at a micro level (i.e., using the idiodynamic method; Nematizadeh & Wood, 2019; Wood, 2016).

Data Elicitation

Before choosing a specific method, the level of potential dynamicity at play for the specific ID, the independent variable of interest, must be considered. For example, does the ID (like WTC or anxiety) change from moment to moment? Or is it more like WM and unlikely to change during an experiment? Given how WTC and state anxiety change frequently, I recommend MacIntyre’s
idiodynamic method. As described in his 2012 article, during a communicative task, participants provide “self-reported ratings on one or more variables of interest” (p. 361). Ratings can be provided for each second of L2 speech, and the data is presented as a graph of changes for each variable during the course of communication. Following task completion, participants (teachers and researchers) are interviewed to explain the observed fluctuations of themselves or their perceptions of others. So, for example, a student participant can explain why their WTC increased strongly after a more fluent partner complimented their work, or how they overcame a high spike in state anxiety and kept going. This method is particularly suited to the most dynamic IDs, including WTC and state anxiety, which have traditionally been measured via questionnaires. When idiodynamic methodology is not possible, I recommend those interested in giving a state anxiety questionnaire, for example, to do so via incorporation in Qualtrics or another electronic option that permits pop-ups during a task and provides less interruption than a paper-and-pencil questionnaire. For WM, I believe there is considerable work to be done to determine which measurements are most appropriate for a given population. Existing research on L2 speaking has often chosen a measure of phonological WM; the measurement chosen should be based on additional learner IDs, especially age and proficiency level. Given the evidence in this chapter that demonstrates an overall positive effect of WM for L2 speaking regardless of measurement, it is beyond the scope here to recommend a specific measurement; the reader is directed to Chapter 3.

Regardless of which ID is of interest, researchers must take care to collect data to assist in their operationalizations of the ID(s) serving as independent variable(s). For example, for age, researchers should collect information about the participant’s age at the time of study, the age of onset at home or in an immersion setting (with additional details on each if possible), as well as the age of onset in a formal schooling setting (with additional details). This will allow for the examination of more subtle relationships between IDs, such as age and context, with the dependent variables of interest for L2 speaking. Even in a study viewing age as binary (pre- or post-puberty, for example), the data is available for additional analyses to provide meaningful insights into these complex, dynamic relationships.

In terms of eliciting samples of L2 speech, the researchers must first specify the goals for the project. For example, is the goal to measure learners’ L2 speech on a given task or a representation of what learners are capable of doing at a certain point in time? Or is the goal more centered on learners’ ability to circumlocute around gaps in their knowledge? For performance on a specific task, one decision to make is deciding between monologic (one-way, participant-only L2 speech) or dialogic tasks (two-way tasks where participants interact with another participant/researcher). Monologic tasks enable a sole focus on learners’ capabilities while controlling the L2 speech they hear and respond to; for these reasons they are also not as ecologically valid. Dialogic tasks are more natural and can elicit a wider range of abilities, but they are also less controllable, dependent upon both participants’ communicative and strategic competence. Tasks and especially task sequences that follow task-based methodology are ideal for keeping a focus on meaning while also being malleable for the desired outcome. For example, researchers can manipulate planning time to increase a focus on fluency or have one task completed in the present tense and the second in the past tense, directing learners’ attention naturally to accuracy (Robinson, 2011). One gap that needs to be addressed is investigation into a potential relationship between learner IDs and spoken functional adequacy, or learners’ ability to complete the task in a way that is pragmatically appropriate (Kuiken & Vedder, 2018). Much like the challenges of multiple measures of WM described above for the independent ID variables, differences in operationalizing CAF (or adding in functional adequacy, CAFFA) abound in the literature. In addition to reading more on these measures (see a recent review in Révész et al., 2016), I recommend that researchers pair measurements of CAF with measures such as learners’ abilities to negotiate for meaning (García-Mayo & Lázaro Ibarrola, 2015), use feedback (Mackey et al., 2010), or engage with the language (Oliver et al., 2017; Svalberg, 2020). Looking at how learners manage L2 speech via interaction provides
an important view into their L2 speech abilities beyond a static product shown by CAF measures. Training and/or triangulating learners’ self-rated abilities (Dolosic et al., 2016) with the teacher, researcher, and/or native speakers would provide a more holistic understanding of how different facets of L2 speech relate to the larger picture, and the ultimate goal of language learning to be communicatively competent L2 users.

Practical Applications

For language teachers tasked with measuring students’ L2 speaking, the most important take-aways from this chapter are as follows. First, the learner IDs outlined here are dynamic, either changing with the passage of time (e.g., age, when conceptualized as years post-birth), or dynamic in terms of being influenced by the L2 learning context (e.g., anxiety and WTC), or learner IDs such as learners’ proficiency level (e.g., WM). Second, while relationships between each ID and L2 speaking were found, there were nuanced differences to the overall positive (WTC and WM) and negative (anxiety and age) relationships between the ID in question and L2 speaking production/development; this is undoubtedly due to the myriad ways both the IDs (in particular, WM) and L2 speaking (CAF) have been measured. Thus, we know IDs influence L2 speaking and must be taken into account when creating assessments.

Before choosing which measurement is best, teachers must decide whether they are interested in learners’ L2 speaking abilities in a given task, or if they are after a more holistic understanding of their learners’ L2 speaking. The more general the L2 speaking abilities targeted, the more measurements will need to be taken and over time (even over one lesson). For example, teachers could have learners’ L2 speaking measured in both spontaneous and practiced ways, manipulating the pre-task planning; they could be measured monologically, responding to a recorded prompt and/or in paired work. Given the known interplay between learner IDs and L2 speaking, we must be cognizant that providing an L2 speaking assessment, even a communicative task, will not test different learners in the same way; teachers would do well to measure and take into account learners’ IDs, particularly WTC and anxiety, to more holistically understand their L2 speaking abilities and ensure that no one type of learner (i.e., more WTC and/or lower anxiety) is outperforming their fellow learners.

In light of the waves of WTC, and resulting in-class participation found by the idiodynamic method, and the fact that learners’ participation increased as they rated their WTC and participation over a class period, I recommend that this method is incorporated into L2 teaching when possible. Echoing larger claims within SLA regarding the importance of learner noticing (Schimdt, 1990) and engagement (Svalberg, 2020), it appears that having learners be involved in tracking their WTC and/or participation increases their L2 production in class. Action research conducted by teachers and students would be of incredible value to see if this pattern holds across contexts and for learners of varying WTC. For example, teachers could have learners interact in dyads and track their reported WTC on questionnaires with their real-time data, and incorporate learners into the analysis stage as both participants (sharing how their own WTC changed throughout the interaction and why) as well as budding linguists (hypothesizing what this means for their L2 learning).

Future Directions

Writing this chapter, it became clear significant work is needed to better understand how IDs relate to L2 speaking. Existing research has found relationships between motivation and interaction in, for example, relation to L2 speaking gains (Hernández, 2010), and the interrelated influence of motivation and/or anxiety on L2 production (Poupore, 2016). Learner creativity has been minimally investigated in relation to L2 speaking (Albert & Kormos, 2004), as has gender (Pica et al., 1991), interlocutor proficiency (Davis, 2009), and interlocutor familiarity (O’Sullivan, 2002), and
not at all in relation to learners’ self-regulation, beliefs, and learning strategies. Even within the four most researched IDs discussed, there is a need to examine the relationship between IDs and L2 speaking in a more nuanced way. For example, are there differences between general L2 anxiety and L2 speaking anxiety in relation to L2 speaking?

The most critical future needs are to incorporate dynamic measurements and clear operationalizations. Research methods must adjust to capture ID dynamicity. While we see this within the idiodynamic method for WTC, other IDs must be researched longitudinally with validated instruments (such as the FLCAS or the numerous tests of WM) to see if and how these IDs change and in what relation (if any) to learners’ L2 speaking. Researchers must also be clear and consistent with operationalizations (for example, regarding the type of motivation they are seeking to measure). Likewise, future studies should focus more on L2 speaking as it develops, in addition to snapshots of L2 speaking abilities, and must do so by selecting specific time windows (seconds or minutes, a task or a class, over two weeks or a semester) with intentionality. To be sure, this chapter opened more questions than it answered. The resounding conclusion is that learner IDs, whether they be cognitive (WM), affective (anxiety), conative (WTC), or sociocultural/demographic (age) in nature, relate to L2 speaking. Considerable work must be done to fully understand their influence on L2 oral production, a skill of central importance for SLA.

Acknowledgments

I am grateful to Daniel Jung, Laura Merino, Madison Wray, and Travis Evans-Sago for their assistance.

Notes

1 Many thanks to Volume Editor Shaofeng Li for his suggestions in describing these two theories.
2 Note that this was opposite from Teimouri et al.’s (2019) meta-analysis, which found L2 decoding (listening) to be much more influenced by anxiety than L2 speaking.
3 The reader is directed to MacIntyre’s website at https://petermacintyre.weebly.com/idiodynamic-software.html for the free software download.

References

IDs for L2 Speaking


Laura Gurzynski-Weiss


