The CATESOL Journal

**MURRAY J. MUNRO**
*Simon Fraser University, Burnaby, British Columbia*

# How Well Can We Predict Second Language Learners' Pronunciation Difficulties?

**Mid-20th–century scholars argued that second language (L2) instruction should be rooted in a comparison of the structural characteristics of the first language (L1) and L2. Their enthusiasm for a "scientific" approach to errors reflected the view, based on the contrastive analysis hypothesis (CAH), that learners' difficulties could be predicted through purely linguistic analyses. Pronunciation seemed particularly amenable to this treatment. If teachers knew their learners' problems in advance, they could presumably design curricula and activities to address their students' needs. Although it soon became clear that many aspects of CAH were seriously flawed, interest in a linguistic account of L2 pronunciation difficulties has persisted. This synthesis of empirical findings from pronunciation research demonstrates that the enthusiasm for error prediction has been misguided, largely because of 2 erroneous beliefs: the *assumption of uniformity* and the *assumption of equal gravity*. The need for an alternative perspective promoting evidence-based teaching practices is demonstrated.**

One of the most frequent requests that my colleagues and I receive from teachers is for a list of English pronunciation errors that learners from particular first language (L1) backgrounds will make. There is no question that the sound system of the L1 exerts a great deal of influence on how second language (L2) pronunciation is learned. It is for that reason that a Mandarin accent sounds different from a Spanish one. The appeal of a predicted error inventory for learners is therefore understandable: Beginning teachers may imagine a scenario in which they know in advance exactly which pronunciation problems will require attention in their classes.

They anticipate using this information to tailor instruction to match students' difficulties so that efficient and effective pronunciation guidance can be provided.

## Error Prediction and Pronunciation Teaching

On the surface, the logic of enhancing teaching by anticipating problem areas seems compelling. In fact, a similar line of argumentation was presented several decades ago when "scientific" comparisons of L1 and L2 were touted as beneficial for curriculum and instruction, whether in grammatical structure or pronunciation. Around that time, Weinreich's (1968) work on language contact (originally published in 1953) had set the stage for important new research on bilingualism. Also, the influence of a behaviorist theory of language (Skinner, 1957) was being strongly felt in pedagogical circles. Among the proponents of linguistic comparisons were Fries (1945) and Lado (1957), who articulated the tenets of the contrastive analysis hypothesis (CAH), a proposal with both theoretical and practical aims. Considerable value was placed on CAH in the behaviorist-inspired audio-lingual method of instruction, which had gained traction in North America and elsewhere. Audio-lingual teaching treated accurate pronunciation as essential to successful L2 acquisition: Learners were required to listen to and mimic native-speaker models so that pronunciation errors could be avoided as much as possible right from the start. Knowing which errors would occur was therefore argued to assist in the design of teaching materials.

As a result of mid-20th–century writing and research on CAH, serious weaknesses in the approach were soon identified. The initial goal of predicting all L2 learner errors on the basis of L1 interference proved unattainable, and a variety of alternative perspectives on errors, such as error analysis (Corder, 1967; James, 1998) were offered and also discovered to be flawed (Schachter, 1974). None of these proposals has enjoyed the level of scholarly attention that was originally accorded to CAH. Rather, pedagogically oriented researchers have generally turned their attention away from linguistically based error accounts. Nonetheless, classroom interest in anticipating students' pronunciation difficulties has persisted to the present time, as evidenced by the ongoing popularity of certain older textbooks, such as Nilsen and Nilsen's (1971) *Pronunciation Contrasts in English*. Its authors claim to pinpoint segmental difficulties for English learners from 47 L1 backgrounds. Nearly 50 years after its initial appearance (and a favorable published review from CAH advocate J. Donald Bowen in 1972), this volume is still in print and ranks highly among Amazon.com's offerings on teaching English pronunciation. Another

text that has remained popular for decades is Swan and Smith's (2001) *Learner English: A Teacher's Guide to Interference and Other Problems* (originally published in 1987), which compares the structural properties of a diverse assortment of languages with those of English.

Irrespective of the longevity of teachers' and students' interests in pronunciation errors, the position I present here is that the practical value of attempting to forecast learners' difficulties before teaching even begins is, at best, minimal. Not only are patterns of pronunciation errors *not* predictable in ways that offer much useful information for teachers, but even if errors could be anticipated, few direct benefits would accrue for the classroom. In the sections that follow, I will explain why the merits of error prediction are greatly overestimated because of two false assumptions: the *assumption of uniformity* and the *assumption of equal gravity*.

## Predictive Failures of CAH

According to CAH, the structural differences between L1 and L2 are the source of learner difficulties. For that reason, CAH scholars, such as Stockwell, Bowen, and Martin (1965), focused their work on classifying the possible types of cross-linguistic differences. Their goal was to develop error hierarchies to capture the varying degrees of difficulty posed by particular L1-L2 mismatches. Such hierarchies were based on comparisons of phonemic categories and allophonic distributions. For instance, a high level of difficulty was proposed for a phonological *split*, a situation in which L2 makes a phonemic distinction between two sounds, where L1 has only a single category in which the sounds are phonologically equivalent. Something akin to this seems to apply to Japanese learners of the English /ɹ/ – /l/ distinction, a notoriously difficult contrast for many of them to acquire. In contrastive terms, Japanese has only a single flap category that subsumes the two English sounds.

When researchers began to test CAH's predicted levels of difficulty, however, their empirical findings often ran counter to expectations. Brière (1966/2017), for instance, created an artificial target sound system (T), consisting of strategically selected phones (16 consonants, 9 vowels) from Arabic, French, and Vietnamese, to be taught to English monolinguals. The T system included items and distinctions predicted by CAH to present varying degrees of difficulty. Although the T inventory is too elaborate to explain here in detail, the following three examples illustrate its differing levels of challenge:

1. *Very challenging.* The phonemic contrast between an aspirated dental stop /t̪ʰ/ and a corresponding unaspirated den-

tal stop /t̪/ in syllable-initial position is predicted by CAH to pose a very high degree of difficulty for English speakers. This distinction is called *divergent*, in the sense that no phonemic contrast exists between such a pair in English, though aspirated and unaspirated stops do occur as allophones in complementary distribution. The initial consonant of *time*, for instance is [tʰ], and many North Americans produce a dentalized unaspirated stop in eighth [eɪt̪θ].

2. *Moderately challenging.* Phonemes requiring *regrouping* of features of L1, in this case a high front rounded vowel /y/ and a high back unrounded vowel /ɯ/, should pose difficulty, but less so than the divergent distinction. English speakers are assumed to know both the front-back feature for tongue position and the feature of lip rounding versus nonrounding. However, the particular *combinations* of features in /y/ and /ɯ/ do not occur in English and must therefore be learned.
3. *Minimally challenging.* A *distributional difference*, in this case the occurrence of the velar nasal /ŋ/ in syllable-initial position, is expected to require some degree of learning on the part of English speakers since this phoneme appears only syllable-finally in L1, as in *ri*n*g* [ɹɪŋ]. However, CAH predicts that it should be noticeably easier to learn the new position than to acquire the other two types of differences because /ŋ/ *does* occur as a phoneme in English.

Brière (1966/2017) instructed his participants on the full T inventory through listening tasks, explicit articulatory descriptions, and production exercises. Subsequently, the learners' productions were recorded and evaluated by native listeners from the relevant backgrounds. While the results confirmed some of the CAH-based predictions, in other cases the outcomes showed marked departures from expectation. For instance, the word-initial stop contrast (i.e., the divergence) proved easy for most learners to produce, despite its absence in English. Also contrary to prediction, the difficulty levels of the two vowels requiring feature regrouping showed a large asymmetry. The front rounded vowel was considerably better produced than its back unrounded counterpart, the latter being produced less accurately than nearly all other sounds in T. Yet nothing from the comparative analysis would appear to predict this difference in difficulty. The paradox was amplified two decades later in a study by Flege (1987), who found that /y/ was produced authentically by adult English learners of French. Contrary to CAH-based expectations, however, Flege found that French /u/, which does have an English counterpart, posed much

more difficulty. (He attributed the results to greater ease for "new" vs. "similar" phones.) Finally, in the Brière (1966/2017) study, the initial /ŋ/ was no better produced than the unaspirated stop, even though /ŋ/ has phonemic status in L1, whereas the unaspirated stop does not. Brière (2017) proposed that existing CAH error hierarchies were based on the wrong level of analysis, arguing that descriptions based only on a phonemic and allophonic breakdown were inadequate, and that "exhaustive information at the phonetic level" would be needed to improve predictive accuracy (p. 91).

**Limitations of Linguistic Analysis**

In many respects, Brière (1966/2017) is an exemplary study from a time when empirical work on pronunciation learning was scant (Strain, 1963). However, in the half century that has passed since then, his recommended "exhaustive phonetic analysis" has never emerged, at least not in a way that is applicable to teaching, nor, on the whole, has linguistically based error prediction advanced much. Rather, it has become clear that the problems with CAH are more far-reaching than its simplistic focus on phonemes and allophones. As early as 1970, for instance, Wardhaugh delivered a sharp objection to the advocates of a "strong" version of CAH, criticizing their apparent belief that:

> …[t]hey should actually be able to carry out their contrastive studies quite far removed from speakers of the two languages, possibly without even knowing anything about the two languages in question except what is recorded in the grammars they are using. (p. 125)

Wardhaugh also commented that a "weaker" version of CAH—one allowing for non-L1–based errors—"has proved to be helpful and undoubtedly will continue to be so as linguistic theory develops" (p. 129). He added, however, that its influence on teaching was likely to wane through time.

In retrospect, the latter prediction has turned out to be true, in part because applied linguists now attend less to purely linguistic influences on L2 acquisition than they did during the heyday of CAH, and they place greater emphasis than before on the array of individual learner factors that affect acquisition of L2 communicative competence. It is widely accepted that cognitive and social factors exert powerful influences on learning that are due to variability in language experience, motivation, and aptitude. Furthermore, ample empirical data now demonstrate large interlearner differences in the production of L2 pronunciation features, even when the learners share the same

L1 background (Derwing & Munro, 2015; Munro, Derwing, & Thomson, 2015). In sum, the major influences on L2 learning extend well beyond purely linguistic matters.

Wardhaugh's (1970) observations touch on a major flaw in CAH and, more broadly, in a school of thought that regards languages themselves as objects of investigation, as though languages could be studied independently of their users. An often unstated aspect of such a perspective is the *assumption of uniformity*: that individual differences in learning trajectories during L2 acquisition are trivial or simply uninteresting. In short, the belief is that if the causes of errors really are *linguistic*, then one should not have to look beyond language structures to understand them.

There is little indication in the research literature that the assumption of uniformity is compatible with pedagogically meaningful insights into phonetic learning. At most, linguistic accounts such as CAH allow generalizations about group performance that do not apply to all learners. A number of recent L2 speech studies, in fact, point to large individual differences in segmental acquisition that undermine the value of group data (Huang & Evanini, 2016; Munro et al., 2015; Nagle, 2018; Smith & Hayes-Harb, 2011). To those interested only in theory building, this limitation might be seen as unimportant because individual variability is regarded as uninteresting noise in the data that falls outside the purview of their abstract accounts (e.g., Gregg, 1990). While a full critique of that particular brand of theorizing is beyond the scope of this article, it should be obvious that any approach to L2 acquisition that allows no explicit mechanism to account for individual variability cannot be of much, if any, use in language teaching. Teachers cannot treat their students as embodiments of a Chomskian ideal speaker-hearer (Chomsky, 1965). Rather, individual variability is not only interesting to teachers, but it is one of the most important aspects of learning that they are required to address in their classrooms.

To understand the practical shortcomings of CAH, and of all other purely linguistic approaches to errors, it is helpful to draw on published L2 pronunciation data. While scores of relevant studies could be referenced here, the sections that follow focus on several aspects of L2 English for which data are particularly informative: dental fricatives and vowel phonemes.

### *Dental Fricatives*

The dental fricative /θ/ is rare as a phoneme in the world's languages, so from a contrastive standpoint, its occurrence in English poses a potential challenge to learners from many L1 backgrounds.

In fact, English learners substitute a variety of phonologically similar segments for /θ/, and the choice of substitute segment has been traditionally argued to depend on the L1.

In a recent study of 36 Mandarin immigrants in the US, Huang and Evanini (2016) evaluated English /θ/ productions in a read-aloud passage. The speakers were all adult arrivals with a mean length of residence of about 10 years. Although Mandarin has no /θ/ category, on 50% of the tokens, the target /θ/ was fully realized. The bulk of the remaining productions were produced as /s/ (29%) or as an interdental stop (12%). Lower frequencies were reported for /sθ/ (4%), /z/ (2%), and /t/ (1%), with a few occurrences of five additional variants. When the interdental stops were classified together with correct /θ/ as "near-native" productions, the authors noted extensive interspeaker variability, with four participants (11% of the cohort) producing 100% of tokens accurately and one scoring 0%. The individual differences were partially explainable in terms of statistically significant relationships between /θ/ accuracy and three speaker variables: initial English exposure, recent English exposure, and language aptitude scores. In short, factors other than L1 clearly influenced performance.

In a different investigation, accuracy of /θ/ was evaluated for Dutch users of English by Wester, Gilbers, and Lowie (2007), who described their 25 participants as a heterogeneous sample of "average" speakers. Like Mandarin, Dutch has no phonemic /θ/. Tokens were obtained from 10-minute extemporaneous English productions. Although the authors did not report the frequency of accurate targets, 64% of non-targetlike productions in word-initial position were /t/, 21% were /s/, and 13% were /f/. In final position, the distributions were strikingly different: 33% were /t/, 13% were /s/, 47% were /f/, and 7% were /d/. Intraspeaker variation was large, with all but two participants showing inconsistent substitution patterns. Once again, the data point to strong influences of factors in addition to L1.

These studies might be seen as shedding light on L2 speaker behavior: They show different performance for speakers of different L1s and effects of position-in-word, and they indicate the phonological features shared by targets and substitutions. However, it must be emphasized that the studies show *tendencies*, rather than consistent performance for individual learners. Consequently, the implications for pedagogy appear to be minimal. Let us suppose that a teacher is assigned to a class consisting entirely of Mandarin or Dutch learners—the simplest-case scenario, but uncommon in any North American ESL context—and undertakes to teach /θ/ through lock-step instruction. On the basis of the Huang and Evanini data, more than 10% of the students (the perfect performers) would presumably be wasting

their time altogether. Moreover, given the overall high level of near-native performance on /θ/, at least a few additional class members would probably show ceiling performance at the outset of instruction with little or no subsequent improvement. On the other hand, some very poor performers might require considerably more instruction than other class members.

Other serious problems for pedagogy are also raised by the data. Since multiple nonnative variants are used by the participants in both studies, it appears that the underlying reasons for the substitutions may differ from learner to learner. In Wester et al. (2007), word-initial substitutions of /t/, /s/, and /f/ might be due to perceptual confusions; however, these confusions apparently differ from speaker to speaker. And the occurrence of /sθ/ in the Huang and Evanini (2016) study may indicate a lack of control over *production* rather than a perceptual difficulty. Given the complexities in the data, conscientious teachers would be faced with the daunting task of keeping track of all the possible confusions in both initial and final positions to ensure effective instruction. But since different learners have different problems, a "one size fits all" strategy for pronunciation teaching is unacceptable. Finally, it is essential to remember that typical ESL classes are not homogeneous with respect to L1. This suggests that interlearner variability would be higher than what is seen in these two studies.

***Vowels***

The concerns raised in the preceding section might be seen as too academic to be taken as a valid criticism of linguistic error prediction. In particular, the study participants were long past the point of being enrolled in regular English classes and had reached a high level of proficiency; their performance might not parallel that of beginner- or intermediate-level students who are more likely to be candidates for pronunciation instruction. It is generally accepted that transfer effects from L1 tend to decline as L2 proficiency increases (Bardovi-Harlig & Sprouse, 2017). To illustrate why this objection is unconvincing, let us turn to data from actual learners in Munro and Derwing's (2008) longitudinal study of Mandarin and Slavic speakers' English vowel productions. At the outset of the study, the participants (*N*s=20 and 24, respectively) were low-proficiency students at an early stage in their first year of ESL classes in Canada. Since the learners' L1s do not distinguish high tense and high lax vowels at the phonemic level, a CAH analysis would predict a tendency to produce /ɪ/ (*bit*) and /ʊ/ (*look*) incorrectly as [i] (*beat*) and [u] (*Luke*), which do exist in the L1s. Findings from Consonant-Vowel-Consonant words produced in a delayed repetition task partially upheld the prediction of difficulty

with /ɪ/: It was the most inaccurately produced of all the English vowels. In fact, only 31% of the Mandarin speakers and 20% of the Slavic speakers were able to produce the vowel in *bit* intelligibly. However, substitutions were not consistently of a single category, and contrary to expectation, the most common substitution was /ɛ/ rather than /i/. To complicate matters, the corresponding lax back /ʊ/ was not especially difficult, with about 70% of speakers in both groups producing *book* with an intelligible vowel. There appears to be nothing in CAH that would account for this discrepancy.

Once again, an attempt to apply CAH in a class of Mandarin or Slavic students would lead to poor instructional choices since the predictions would be mostly wrong, even in terms of general tendencies. But a more serious problem would arise from the dramatic individual differences in performance. Some learners produced all high vowels very accurately right from the start, and some produced most of the other vowels accurately as well. On the other hand, sizeable numbers—though not everyone—could probably have benefited from focused instruction on the /ɪ/-/ɛ/ distinction, while perhaps a minority could improve their productions of /ʊ/. These concerns cannot be dismissed as too hypothetical, since the students were in fact enrolled as ESL students at the time of data collection. Nor is this study the only one to report large-scale interspeaker variability in vowels produced by speakers sharing an L1. Rather, it mirrors an earlier study of Italian speakers in Canada (Munro, Flege, & MacKay, 1996) in which the authors concluded that generalizations of the type "Speakers of Italian produce vowel *x*, but not vowel *y*, in a native-like way cannot be made here because the between-vowel effects did not occur uniformly for all, or even for a large majority of learners" (p. 332).

The uniformity assumption is not confined to CAH. Rather, it is a fundamental problem with *any* approach to the study of L2 errors that treats them in purely linguistic terms. The markedness differential hypothesis (Eckman, 1977), for instance, was intended as a refinement to CAH. In Eckman's approach, the degree of difficulty posed by L2 structures (phonological and other) was to be predicted through a comparison of L1 and L2, with the incorporation of typological markedness as an additional linguistic influence. While differences between L1 and L2 were still assumed to contribute to errors, more marked features of L2 were expected to pose greater difficulty than less marked features. For instance, the voiced/voiceless distinction is more marked (cross-linguistically less common) in word-final than in word-initial position. This fact allegedly accounts for why German speakers find it difficult to distinguish word pairs such as *pick/pig*. Whatever the theoretical merits of this approach, it includes no

mechanism for addressing individual differences and therefore may have no more applicability to teaching than does an unmodified version of CAH.

Still one other linguistic approach is the application of optimality theory to explain L2 segmental errors in terms of abstract constraint rankings (Wester et al., 2007). For instance, differences in L1 rankings are supposed to account for why Russian speakers are said to realize English /θ/ as [t], while Japanese speakers are claimed to substitute [s] (Lombardi, 2003). In fact, Lombardi offered the flat assertion that "[s]ubstitutions for English interdentals tend to be consistent based on first language" (p. 225) and proceeded to offer an intricate argument for why Russians realize /θ/ as [t]. Yet her work presents no data whatsoever to support her claim of consistency. More recent work by Munro et al. (2015) showed that 83% of /θ/ productions by beginning-level Slavic speakers were actually on target: Few of the learners actually had difficulty with the sound. Moreover, though not specifically reported in the paper, several of the incorrect substitutions were [s]. In fairness to Lombardi (2003), her analysis was not intended to be relevant to the classroom, but ironically the Munro et al. (2015) data call into question the theoretical value of her complex explanation for a phenomenon that may not even exist.

## Error Prediction and Error Gravity

The points raised above are clearly damaging to the claim that error prediction is valuable in pronunciation teaching. However, there is another fundamental problem with the concept, namely, the *assumption of equal gravity*. Here I am referring to the belief that by identifying errors in L2 production we have simultaneously identified a suitable focus of instruction. In that view, any error is just as deserving of a teacher's attention as any other. This impression is conveyed, perhaps unintentionally, by such texts as Nilsen and Nilsen (1971), which simply list predicted segmental difficulties with no indication that some might be more important than others and no explanation of how their significance compares with that of suprasegmentals. The assumption of equal gravity has been thoroughly discredited in contemporary perspectives on pronunciation teaching (Derwing & Munro, 2015; Levis & Cortes, 2008). Some errors may result in serious communication breakdowns because of a loss of intelligibility, while others may be noticed as an aspect of a speaker's accent, yet have no detrimental effect on the listener's success in processing the speech. The above discussion of English learners' difficulties with /θ/ and /ɪ/ provides an example of just such a difference. While both segments appear to pose considerable difficulty for many learners, inaccurate

production of /θ/ seems unlikely to create communication difficulties. In the first place, Catford (1987) reported that English contrasts with /θ/ score relatively low in terms of functional load, a concept referring to the "amount of work" performed by sound pairs in keeping words distinct. For instance, the English /p/-/b/ opposition has a high functional load since many minimal pairs depend on it (e.g., *pat/bat*, *pin/bin*, *pay/bay*, etc.). Moreover, /p/ and /b/ occur in many high-frequency words and in words that might be confused with each other because they represent the same part of speech (e.g., *pill* and *bill* are both nouns). In contrast, few confusable minimal pairs exist for /θ/-/s/, /θ/-/t/, and /θ/-/f/. In the second place, Munro and Derwing (2006) found evidence that errors involving /θ/ (e.g., using /f/ instead) had only minimal effects on speech comprehensibility. When deciding where to focus their attention during instruction, teachers should note that /i/-/ɪ/ and /ɛ/-/ɪ/ distinguish very large numbers of minimal pairs in English and therefore have high functional loads. One might therefore expect that incorrect productions of /ɪ/ should have more serious implications for speakers than should incorrect /θ/ (see Levis & Cortes, 2008).

## Conclusions

I began this article by referencing a common request from teachers for lists of pronunciation errors likely to be made by speakers of different L1s. In fact, linguistically based lists of this type have existed for decades (Nilsen & Nilsen, 1971; Swan & Smith, 2001) and have enjoyed remarkable popularity. The available evidence, however, indicates that they are minimally useful to teachers. The following underlying problems have been identified:

1. At best, linguistic analysis can yield accurate error predictions only in terms of general tendencies. These have little classroom applicability, first because even the predicted tendencies are often wrong, and second because of the false *assumption of uniformity* across learners. Moreover, there is no reason to think that linguistic error prediction can ever be much more accurate than it already is. This is because it fails to take into account the array of nonlinguistic factors (including individual differences in experience, aptitude, and motivation) that influence L2 phonetic learning.
2. Focusing on error prediction tends to be linked with the false *assumption of equal gravity*, the view that identifying errors is equivalent to identifying foci for teaching. It is not true, however, that every error that emerges in L2 speech is equally in

need of correction. Ample evidence indicates that some errors are far more detrimental to communication than others. For that reason, even successful error prediction gives teachers no help in deciding which aspects of pronunciation they should actually emphasize in their classes.

To develop a full understanding of how phonetic learning takes place, a consideration of errors seems essential. However, attending only to linguistic influences on acquisition cannot advance knowledge about L2 pronunciation beyond what is already known. In that regard, promising developments in the field include Best's perceptual assimilation model (Best & Tyler, 2007) and Flege's speech learning model (Flege, 1995), both of which devote attention to the *perception* of non-L1 speech sounds. Both also diverge from traditional linguistic approaches such as CAH in that they do not treat languages as the objects of study. Flege's model, in particular, assumes that as learning progresses, some aspects of the L2 system can be acquired at a perceptual level and that production will eventually align with the newly developed perceptual knowledge. To account for why one learner produces a particular phone correctly and another from the same L1 background does not, one might well look at differences in perception. These may arise from or interact with influences such as L2 experience and L1 use (Piske, MacKay, & Flege, 2001). Nonetheless, such models suggest that the variables that determine a particular learner's success with a particular sound are too complex to be assessed and then used in classroom settings.

How well can we predict L2 learners' pronunciation difficulties? Not well enough to offer anything more than weakly relevant information to teachers. This is true enough when classrooms are homogeneous with respect to L1, but when classes are linguistically diverse, the value of error prediction decreases still further. At best, one can point out a handful of difficulties that some (occasionally a majority) of learners might have at the outset of learning, though some of these may not actually merit instruction. It follows that instead of focusing their attention on predicting errors, teachers should invest their energies in assessing the actual needs of individual learners and helping them to address their most communicatively important difficulties. Fortunately, the growing body of work on computer-assisted pronunciation teaching suggests that individualized instruction will become more available and more effective in the future (Munro & Derwing, 2015). As the field of pronunciation teaching moves forward, our best strategy is to leave our fascination with error prediction behind.

**Acknowledgment**

I am indebted to Tracey Derwing, who shared in the development of many of the ideas in this manuscript. All errors are entirely my own.

**Author**

*Murray J. Munro is a professor of Linguistics at Simon Fraser University and co-editor of the* Canadian Modern Language Review. *He and his colleague Tracey Derwing have published extensively on L2 pronunciation, including their book* Pronunciation Fundamentals. *His most recent volume is* Critical Concepts in Pronunciation, *co-edited with John Levis.*

**References**

Bardovi-Harlig, K., & Sprouse, R. A. (2017). Negative and positive transfer. In J. Liontas (Ed.), *The TESOL encyclopedia of English language teaching*. Hoboken, NJ: Wiley. doi.org/10.1002/9781118784235.eelt0084

Best, C., & Tyler M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Amsterdam, The Netherlands: Benjamins.

Bowen, J. D. (1972). Review of "Pronunciation contrasts in English" by D. L. F. Nilsen & A. P. Nilsen. *The Modern Language Journal*, *56*(3), 190-191.

Brière, E. J. (2017). An investigation of phonological interference. In J. Levis & M. J. Munro (Eds.), *Pronunciation* (pp. 61-94). London: England: Routledge. (Original work published 1966 in *Language*, *41*[4], 768-796)

Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 87-100). Washington, DC: Teachers of English to Speakers of Other Languages.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, *5*(1-4), 161-170.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* Amsterdam, The Netherlands: Benjamins.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*(2), 315-330.
Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65.
Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233-277). Timonium, MD: York Press.
Fries, C. C. (1945). *Teaching and learning English as a foreign language.* Ann Arbor: University of Michigan Press.
Gregg, K. R. (1990). The variable competence model of second language acquisition, and why it isn't. *Applied Linguistics*, *11*(4), 364-383.
Huang, B., & Evanini, K. (2016). Think, sink, and beyond. *Journal of Second Language Pronunciation*, *2*(2), 253-275.
James, C. (1998). *Errors in language learning and use: Exploring error analysis.* London, England: Longman.
Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.
Levis, J., & Cortes, V. (2008). Minimal pairs in spoken corpora: Implications for pronunciation assessment and teaching. In C. A. Chapelle, Y.-R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 197-208). Ames: Iowa State University.
Lombardi, L. (2003). Second language data and constraints on manner: Explaining substitutions for the English interdentals. *Second Language Research*, *19*(3), 225-250.
Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*(4), 520-531.
Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, *58*(3), 479-502.
Munro, M. J., & Derwing, T. M. (2015). A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation, 1*, 11-42.
Munro, M. J., Derwing, T. M., & Thomson, R. I. (2015). Setting segmental priorities for English learners: Evidence from a longitudinal study. *International Review of Applied Linguistics in Language Teaching*, *53*, 39-60.
Munro, M. J., Flege, J. E., & MacKay, I. R. (1996). The effects of age of

second language learning on the production of English vowels. *Applied Psycholinguistics*, *17*(3), 313-334.
Nagle, C. (2018). Motivation, comprehensibility, and accentedness in L2 Spanish: Investigating motivation as a time-varying predictor of pronunciation development. *Modern Language Journal*, *102*(1), 199-217.
Nilsen, D. L. F., & Nilsen, A. P. (1971). *Pronunciation contrasts in English*. New York, NY: Simon and Schuster.
Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191-215.
Schachter, J. (1974). An error in error analysis. *Language Learning, 24(2)*, 205-214.
Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley.
Smith, B. L., & Hayes-Harb, R. (2011). Individual differences in the perception of final consonant voicing among native and non-native speakers of English. *Journal of Phonetics*, *39*(1), 115-120.
Strain, J. E. (1963). Difficulties in measuring pronunciation improvement. *Language Learning*, *13*(3-4), 217-224.
Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1965). *The grammatical structures of English and Spanish*. Chicago, IL: University of Chicago Press.
Swan, M., & Smith, B. (Eds.). (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge, England: Cambridge University Press.
Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly*, *4*(2), 123-130.
Weinreich, U. (1968). *Languages in contact* (2nd ed.). The Hague, The Netherlands: Mouton.
Wester, F., Gilbers, D., & Lowie, W. (2007). Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*(2), 477-491.