1 Introduction

Phonology is the study of sounds in languages. It is the field that is concerned with questions about which sounds in languages are used to contrast meaning, how sounds pattern together in terms of phonotactics (which sequences of sounds are permissible) and which phonological processes (such as assimilations or elisions) take place. Second language (L2) phonology, then, is the field that investigates how language users acquire the sound system of a language that is not their native or home language. The acquisition process in L2 phonology is fundamentally different from that in first language (L1) phonology, as L2 learners have already acquired the phonology of their native language, but now need to acquire a second sound system that is different from the one they had learnt from birth. Research has firmly established that this is not a trivial task: on a segmental level, for instance, many L2 learners struggle with the acquisition of sounds which are not contrastive in their native language. Moreover, prosodic differences, such as differences in rhythm, may pose additional challenges to the learner.

This Element focuses on the acquisition of L2 phonology at a segmental level, meaning at the level of consonants and vowels, and specifically discusses the impact of phonetic variation on the acquisition of an L2 phonological system. When native speakers of Dutch, for instance, are acquiring the sound system of English, they need to learn that the sounds /ɛ/ (as in 'bed'), and /æ/ (as in 'bad') are contrastive. As Dutch has only one vowel in this area of the vowel space, even the discrimination between these two English sounds and the perception of the sounds as distinct categories turn out to be challenges. Only with sufficient high-quality exposure may L2 learners succeed in building robust phonological representations in their L2. In addition, English sounds may be realized differently by different speakers of English, depending on, amongst other factors, the regional background of the speaker. In Australian English, for instance, the difference between the vowels /ɪ/ ('ship') and /i/ ('sheep') is spectrally reduced (Cox & Palethorpe, 2007). When L2 learners who have mostly been exposed to, for instance, American English, are then listening to an Australian English speaker, they need to adapt their perception to the accent of the speaker. Another example is the word 'better' produced in three different ways: with a plosive [t] and a final schwa by a speaker of British English (Speech sample 1), with a flap and a final [ɐ] by an Australian English speaker (Speech sample 2) and with a glottal stop and a ɐ, again by a speaker of Australian English (Speech sample 3). As shown in Figure 1, these three realizations of the same word differ considerably in terms of the acoustics of both the medial consonant and the final vowel. Despite this variability in the acoustic signal, listeners need to be able to map all three

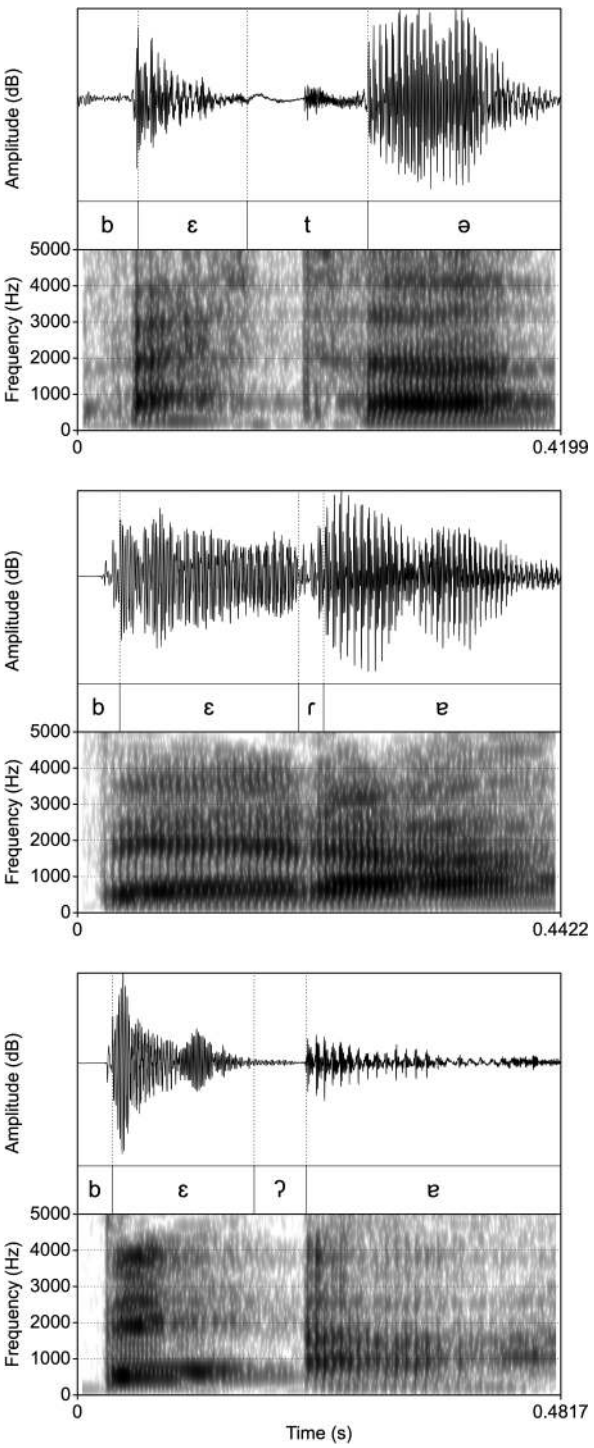

Figure 1 Different realizations of the word 'better' by a British English (top) and Australian English speakers (middle and bottom).

sounds to the same lexical item, 'better'. This is known as the 'lack of invariance' problem. This term was coined by Liberman et al. (1967) to refer to a long-standing issue in speech perception research which illustrates the complex relationship between acoustic cues and phonemes.

Speech sample 1 'better' with a plosive – British English speaker. Audio file is also available at www.cambridge.org/Simon
Speech sample 2 'better' with a flap – Australian English speaker. Audio file is also available at www.cambridge.org/Simon
Speech sample 3 'better' with a glottal stop – Australian English speaker. Audio file is also available at www.cambridge.org/Simon

This means that the learners' categories need to be sufficiently flexible to deal with all the variation that learners encounter. It is inherent that, whenever language learners come into contact with their target language, they are also automatically exposed to variation, including socio-indexical variation, which provides listeners with information on the regional and social background of the speaker. Despite this observation, relatively little research so far has addressed whether or to what extent L2 learners acquire sociophonetic information in an L2. As we will argue, the study of sociophonetic variation can enhance our understanding of how language users build phonological representations. Hence, insights from sociolinguistics and particularly from sociophonetics are relevant to the development of phonological theory. In addition, studies in clinical linguistics that focus on the phonology of atypically developing language users may provide further insight into the nature of phonological representations. If we come to understand whether – and if so, how – reduced language skills in this population can be related to phonological representations, insights from clinical linguistics may tell us something about how phonological representations are organized in typically developing L1 and L2 learners.

The Element centres around the following major topics:

- The field of L2 phonology (Section 2)
- The development of robust phonological representations in an L2 (Section 3)
- The flexibility of phonological representations (Section 4)
- The link between sociophonetics and L2 phonology (Section 5)
- Fuzzy and overspecified representations (Section 6)
- Assessing the formation of categories (Section 7)

In Section 8, avenues for future research will be explored as new ways to advance the field of L2 phonology in terms of methodology and theory. Conclusions are formulated in Section 9.

2 The Field of Second Language Phonology

2.1 A Field at the Crossroads of Different Disciplines

As noted in the introduction, scholars in the field of L2 phonology examine how language users acquire a phonological system in a language that is not their native or home language. The study of this acquisition process touches upon elements from different linguistic subdisciplines and requires insights from different domains, including L2 acquisition, phonological theory, phonetics and sociolinguistics, as visualized in Figure 2.

First, L2 phonology research is embedded in the field of L2 acquisition. In any type of language acquisition, three main types of 'actors' play a role: the learner, the context and the interlocutor or (target) recipient of the message that the learner wants to get across. A myriad of factors related to the learner impact the L2 learning process, including the learner's native language, age, age of first exposure, aptitude, motivation, length of residence in the country/region of the target language, amount of L2 use and socio-economic and educational background. The relative impact of these factors will depend on the context in which the learning process takes place. Important factors related to the context are the naturalistic versus instructed setting in which learning takes place, and the amount and quality of exposure, including the number of different interlocutors. Finally, there is a growing body of research on native listeners' perception of (L2) accented speech, comprehensibility (listeners' self-reported level of

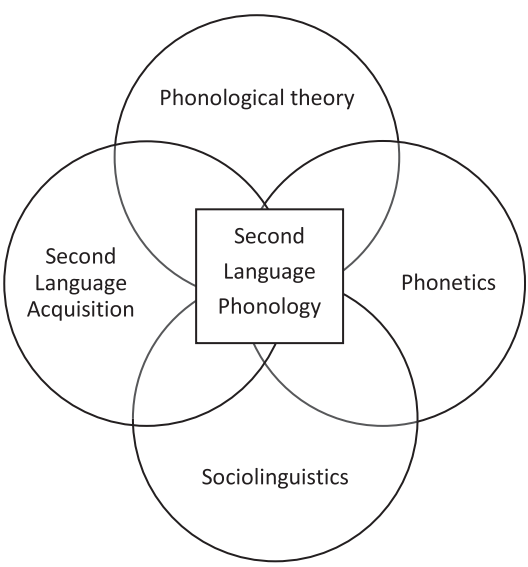

Figure 2 Second language phonology at the intersection of different disciplines.

understanding) and intelligibility (speech actually understood by the listeners). The speaker's pronunciation (segmental properties, prosody and fluency) and lexicogrammar (richness of vocabulary, as well as accuracy and complexity of grammar) have been shown to be the two main linguistic dimensions influencing a speaker's intelligibility and comprehensibility (Saito, Trofimovich & Isaacs, 2017). However, research on L2 phonology has increasingly recognized the impact of the interlocutor. Crucially, a learner's L2 speech can only be called (in)comprehensible or (un)intelligible *to a certain listener* (Munro, 2008; Simon, Lybaert & Plevoets, 2022). As a result, it is essential to recognize the role of the listeners in L2 phonological acquisition, including listener characteristics related to, for instance, native language, age, familiarity with L2 speech, social and sociopolitical attitudes, socio-economic status and educational background.

Secondly, in order to address issues in L2 phonology, questions that are fundamental to phonological theory need to be raised. Such questions can be related to the nature of phonological units, the way these units are organized and their level of abstraction. In order to build models of L2 phonological acquisition, the issue of what exactly needs to be learnt by the L2 speaker-listener needs to be addressed. When L1 learners start building phonological representations, they do so from scratch ('bottom up'), on the basis of phonetic input. Conversely, adult L2 leaners have a full set of L1 representations available, including their phonetic implementations. As a result, they may additionally use a 'top-down' approach in their processing of an L2, looking for correspondences between speech elements in their L1 and in the L2 (Flege, Munro & MacKay, 1995: 22).

Thirdly, research in L2 phonology draws on the field of phonetics, which is the science of how speech sounds are articulated, how they transfer through the air in the form of waveforms and what acoustic properties they have, and how they are perceived by the listener. Most current L2 speech acquisition studies start from the acoustic input: learners acquire an L2 sound system by being exposed to acoustic-phonetic forms (Flege & Bohn, 2021). As such, work in L2 phonology will necessarily draw on insights in the closely related discipline of phonetics. Much of the earlier work in laboratory phonology was devoted to the question of how phonology is related to phonetics (Cohn, Fougeron & Huffman, 2017) and this relationship remains a crucial factor in current discussions of L2 phonological research.

Finally, as variation is an inherent property of speech, one of the challenges faced by L2 learners is learning how to deal with this variation, which is at the same time linguistic and socio-indexical in nature. For instance, a vowel in a particular language may have different phonetic realizations, from more open

to more closed and from more front to more back, depending on the flanking consonants (e.g. more fronted preceding a front consonant), but also on the speakers' regional or social backgrounds. When studying patterns of L2 phonological acquisition, we therefore also need to take into account insights from sociolinguistics, or – more specifically – from sociophonetics.

In sum, the field of L2 phonology is a discipline situated at the interface of research on L2 acquisition, phonology, phonetics and sociolinguistics. Non-native listeners need to build robust phonological representations in their L2, which can be impeded by their native language phonology (Section 3), different listening conditions (Section 4) or sociophonetic variation (Section 5). We will explore these challenges to L2 learning and zoom in on a group of learners that may help us to understand the nature and development of L2 phonological representations, namely child and adult atypically developing language users, who have been claimed to have fuzzy representations in their L1 (Section 6). In the next section (Section 2.2), we go back to a basic question in theoretical phonology that is key to the field of L2 phonology: what needs to be learnt?

2.2 What Needs to Be Learnt?

When people learn a language that is not their native language, they have to learn a new sound system. Even when the language they are learning is similar to their native language(s) in terms of, for instance, the grammar or the lexicon, there will undoubtedly be differences in the way in which segments are produced and interact with each other. The target language may have a sound that does not occur in the native language or not in the same positions (i.e. phonotactics may be different), it may have connected speech processes, such as assimilation patterns, which do not apply in the native language, or it may differ in the concrete realization of sounds. Hence, the question emerges what exactly needs to be learnt by learners who need to master the sound system of an L2. As Flege and Bohn (2021) point out, earlier speech learning models, such as contrastive analysis, made predictions regarding areas of difficulty in L2 acquisition, which were based on comparisons of the sound inventories at the *phonological* level: phonemes which did not occur in the L1 of the learner were predicted to be difficult. As these predictions were often not borne out, the idea that the *phonetic* level needs to be taken into account gained ground in the field of L2 speech acquisition research. Indeed, most current L2 acquisition studies now start from the acoustic input: learners are exposed – either in naturalistic or instructed settings – to acoustic-phonetic forms and acquire phonetic categories on the basis of this input. Most prevalent speech learning models – including the Speech Learning Model (SLM, Flege, 1995) and its revised version (SLM-r, Flege & Bohn, 2021), the Perceptual Assimilation Model (PAM, Best, 1995) and its adaptation for

L2 speech learning (PAM-L2, Best & Tyler, 2007) and the L2 Linguistic Perception Model (L2LP, Escudero, 2005) and its revised version (van Leussen & Escudero, 2015) – aim to account for how learners develop categories for L2 sounds on the basis of the phonetic input. In fact, in L2LP, it is even argued that a detailed acoustic comparison of L1 and L2 sounds can reliably predict L2 perception, since the acoustic properties of listeners' L1 speech sounds will be the starting point of initial L2 perception (Elvin, Escudero & Vasiliev, 2014). A useful overview of current speech learning and perception models including PAM, SLM, the Native Language Magnet Model and Automatic Selective Perception is provided by Chang (2019). Baese-Berk et al. (2022), focusing on perception, point out that most speech perception models rely on 'category learning'. They define a category as 'an abstract and generalizable representation that enables listeners to perceive highly variable acoustic input and efficiently process it through a more parsimonious representation with fewer perceptual dimensions than the raw sensory input' (3026).

Indeed, learners are exposed to a great deal of phonetic variation in the input and thus need to build representations at a more abstract level. In the next sections we aim to present a coherent overview of how this process takes place, how representations develop during the acquisition process and when the process may be hampered.

3 Towards Robust Phonological Representations in a Second Language

3.1 Developing Phonological Representations in First and Second Languages

Before we turn to the question of how L2 learners acquire L2 categories, it is interesting to consider how the process unfolds in L1 acquisition. After all, when adult language users learn an L2, the phonological system of their L1 will to a great extent determine the initial state of the L2 acquisition process.

Research on L1 acquisition has firmly established that after the first six months of life, children start to develop language-specific categories as a result of perceptual attunement to the native language (Kuhl, 1992; Polka & Werker, 1994; Werker & Tees, 1984; for a review, see Werker, 2018). This attunement is assumed to be the result of an increased knowledge of the phonological status of relevant L1 distinctions. In Kuhl's Perceptual Magnet Model (Kuhl, 1992), infants deduce this information through exploiting statistical properties, in the sense that highly frequent and hence familiar speech sounds would function as magnets by attracting similar sounds and thereby diminishing the discrimination between these similar sounds (Werker, 2018: 710). Throughout childhood and adolescence, speakers' phonological representations keep developing: as children grow older

and turn into adolescents, the way in which they use acoustic cues while listening becomes more adult-like and perception gradually becomes more categorical. By the time they reach adulthood, they have normally built well-established, robust categories in their L1.

When adults then acquire an L2 they need to develop a new set of categories for the L2. A large body of research on L2 perception and word recognition has amply demonstrated that this is not a trivial task. In fact, acquiring new categories for contrasts which are absent in the L1 is known to be very difficult, even for highly proficient L2 learners. Examples are studies on the /r/-/l/ contrast in L2 English for L1 Chinese listeners (Aoyama et al., 2004; Cutler & Otake, 2004) or the /ɛ/-/æ/ contrast in L2 English for L1 Dutch learners (Broersma, 2005; Escudero, Simon & Mitterer, 2012; Simon, Sjerps & Fikkert, 2014). Previous research has, however, also established that L2 learners may be able to create new categories by shifting the boundaries of L1 categories in the direction of the L2 (Elman, Diehl & Buchwald, 1977; Flege & Eefting, 1987). According to models such as SLM-r and PAM-L2 (see Section 1), the likelihood of successful category creation is predicted to depend on the L2 category's relation to existing L1 categories. In SLM-r, the L2 sound's degree of perceived phonetic dissimilarity from the closest L2 sound will determine whether a new category is formed or not, in addition to the quantity and quality of L2 input and 'the precision with which the closest L1 category is specified when L2 learning begins' (Flege & Bohn, 2021: 65). In the framework, category precision is defined by the degree of acoustic variability that is produced by a speaker in multiple productions of the category. The authors refer to a study by Kartushina and Frauenfelder (2013), which showed that Spanish learners of French whose L1 Spanish /e/ productions were 'compact', in the sense of revealing relatively little token-to-token variability, were better at identifying French /ɛ/ than Spanish speakers with a less precise Spanish /e/ category – that is, in which there were more spectral differences between different /e/ realizations. Flege and Bohn (2021) point out that category precision is linked to the distance between the category and categories which are adjacent in phonetic space, but may also vary between different speakers, possibly as the result of differences in, for instance, auditory acuity and working memory (39).

As pointed out by Chang (2019), the L1 and L2 are linked not only at a segmental level, as predicted by SLM-r, but also at a broader, systemic level. Evidence for this comes from studies demonstrating the impact that an L2 can have on features of the L1, a process typically referred to as 'phonetic drift' (see also Section 8.2). Phonetic drift has been found in both immersion contexts (e.g. Chang, 2012, 2013 on English–Korean interactions in L1 English participants immersed in the L2 in South Korea) and in L1-dominant