Listening in interaction: Reconceptualising a core skill

Jonathon Ryan

For many language learners, listening represents a major source of anguish, with apparent success in the confines of the classroom failing to be mirrored in the ordinary interactions of daily life. One contributing factor may be the continued reliance on listening texts and activities that position the learner as a bystander rather than a participant. In response, the concept of *interactive listening* has drawn considerable attention in ELT. What has been largely missing, however, is the application of principles drawn from the last 20 years or so of empirical research in fields such as conversation analysis which reveal how interactions actually unfold for listeners. This article identifies five core concepts that may be fundamental to further developing a robust approach to *listening in interaction* (L-in-I): participation status, meaning as action, projection, participatory listenership, and recipient design.

Introduction

Of the four main language skills, many learners report having the greatest difficulty with listening. Despite this, listening is the least researched and theorized of the four (Vandergrift 2007), is 'arguably the least understood and most overlooked' by teachers (Nation & Newton 2009: 37) and is characterized by the least developed and most routinized pedagogy (Siegel 2015). There are, then, strong grounds for re-considering this pedagogy, looking towards either incremental changes or more fundamental shifts in practice. This paper outlines the case for further developments within what has become known as *interactive listening*, based on taking seriously the nature of listening as it occurs in interaction. Alongside restating the important principles of participation status and listener response, three further concepts are introduced from the conversation analysis literature, each of which has been largely overlooked in ELT listening: meaning as action, prospective meaning, and recipient design. These five principles are likely to prove foundational to developing a robust pedagogical approach that prepares learners for the demands of listening in everyday situations.

The starting point for this paper is recognition that the most elemental site of listening is in interactions. These include such contexts as casual conversations, service encounters, discussions and most workplace talk, whether face-to-face or distance. Central to the character of interaction is participants alternating between turns at listening and turns at speaking, with occasional overlap between the two. As should become clear, such listening has a nature and complexity that suggests a need to develop pedagogical approaches beyond those appropriate to unidirectional texts (e.g. listening to television). In response, various proposals have been made under the

umbrella label *interactive listening* (e.g. Field 2008; Vandergrift 1997), though activities in published textbooks remain overwhelmingly unidirectional.

Though wary of proliferating terminology, it may be useful to distinguish two poles within a continuum of approaches to interactive listening. The first relates to classroom activities based on creating opportunities for students to practice listening as part of a conversation or other types of performance (e.g. Vandergrift 1997). This typically involves an information exchange, with listeners often negotiating and responding to the information their partner provides. Such procedures sit comfortably within the framework of *natural approaches*, in which second language acquisition is promoted through meaningful communication. The second, for which the term *listening-in-interaction* (L-in-I) is introduced, adopts an overtly analytical stance, in which elements of the schematic knowledge that interactants draw upon when listening are identified, and then specifically – even explicitly – targeted. This is closer in emphasis to suggestions offered by Field (2008: 69-73), who proposes various aspects of modelling and task types, though it seeks to extend that work through underscoring empirical insights into how interactions unfold from a listener perspective.

Participation framework

As long recognized, when working with listening texts, learners are almost exclusively positioned as eavesdroppers rather than as interactants. However, there has been relatively little consideration of the implications of such positioning and what it implies for teaching. A useful starting point is Goffman's (1981) *participation framework*, which acknowledges that anyone within hearing distance of an interaction has a distinct 'participation status relative to it' (1981: 3). Goffman draws a primary distinction between *ratified* participants, those who are officially part of the interaction, and *non-ratified* participants, those who are merely in the vicinity. These statuses entail different listening behaviors.

Participation framework
    Ratified listeners
        Addressees
        Auditors
    Non-ratified listeners
        Overhearers
        Eavesdroppers

(adapted from Goffman 1981)

The *addressee* is the listener (or listeners) to whom an utterance is ostensibly directed. Outwardly at least, it is for addressees that utterances are specifically designed to be understood (Sacks & Schegloff 2007), and this may influence the topic, register, speed of delivery, degree of politeness and so on. Crucially, not only do addressees have certain rights to respond, but they are even obliged to do so, such as in providing appropriate response tokens such as *Mm* and *Oh* (Gardner 2001). In multiparty talk, *auditors* are third parties not currently being addressed. They often remain alert to what is being communicated and display attention through brief responses. They are available as a potential next addressee and to potentially take the next turn, especially as topics end

and speakership lapses, or where the communication is relevant to them. Though they are not being addressed, speakers display sensitivity to auditors through language choice and topic selection.

In contrast, contributions by mere bystanders (even as small as *Mhm*) are almost always unwelcome. Bystanders often pretend to be oblivious to talk by enacting 'a show of disinterest' and often physically withdraw from the audible space (Goffman 1981: 132). On the rare occasions when bystanders do intervene, they are likely to preface contributions with an apology or explanation (e.g. *I couldn't help but overhear …*). There are in fact two socially formulated classes of non-ratified listening: bystanders count as *overhearers* when they inadvertently catch brief stretches of talk and count as *eavesdroppers* when they listen surreptitiously and with intent (Goffman 1981: 131-132). Superficially at least, in listening classes students are routinely positioned in ways consistent with eavesdropping: intently focused on audio texts involving remote speakers.

The question that should concern us is whether the orthodoxy of eavesdropping might shape or constrain the types of listening practice and skill development that language learners receive. In other words, if learners almost always engage with listening texts from a non-ratified positioning, are they missing out on instruction and practice of crucial sub-skills? As discussed below, listening as an addressee requires a substantially more complex skillset than is required by overhearers, yet these skills are almost never addressed in published listening materials.

> **Implication:** learners should have the opportunity to work with listening texts while positioned as addressees and auditors.

Meaning as action

As Field (2008) details, listening pedagogy has long been dominated by a *comprehension approach* in which answers to post-listening questions are treated as evidence of understanding and as the basis for correction. Such questions are overwhelmingly concerned with *referential* meaning in the broad sense of providing information on who, what, when, where, why, and how.

Although referential meaning remains highly relevant to L-in-I, so too is another, arguably more fundamental aspect of meaning that is almost entirely absent from the ELT literature. This involves interpreting the *action* that the speaker is performing. Consider the following scenarios performed in an unfamiliar language:

1. A neighbor cheerfully calls out in passing
2. A street performer beckons you as you walk by
3. An enraged football fan eyeballs you and yells

Whatever the precise wordings and subtleties of meaning, most important is what you determine the speaker to be *doing* and what you should do in response. These particular scenarios involve isolated actions that align easily with one or more familiar scenarios, but in the flow of ordinary talk a great number and variety of actions may

occur in succession and a speaker may perform multiple actions within the same turn (Enfield & Sidnell 2017). As one turn follows another in discourse, and meanings and actions compound, determining what the speaker is doing often becomes vastly more complex.

As Enfield and Sidnell (ibid.: 130) note, this view of communication as *action* is well established within the philosophy of language and is foundational to conversation analysis. As they further state (ibid.: 130):

> 'by speaking we are not simply, solely, or primarily engaged in describing the world, depicting it, or indexing it in some way. Rather, by speaking we are acting in it. When you say *That's a really nice jacket*, you've not only described someone's clothing but also given them a compliment'

Moreover, a compliment is not always *merely* a compliment but may represent a further level of action. Perhaps the speaker is merely being sociable, but among other things could be asserting power, flirting, angling towards borrowing the jacket or discussing this winter's fashion[2]. Having competent skills for L-in-I enables us to respond in the ways that we would want.

At this point, it is worth warding off a potential misapprehension: it is certainly *not* the case that learners should ordinarily be listening to texts and labelling actions (compliment, apology and so on). Even for the expert language analyst, most attempts at labelling are doomed to failure: there is no master list of actions and although a few actions are so familiar as to have a label, most defy ready description (Enfield & Sidnell 2017). What counts is that competent listeners are masters at interpreting utterances in such a way that they know how to respond.

> **Implication:** L-in-I activities should include a focus on speaker action and listener response.

Retrospective vs. prospective meaning

For interactions to succeed, participants must monitor both what has been said (the *retrospective* view) and where the talk may be heading (the *prospective* view) (Deppermann & Günthner 2015). The retrospective orientation exclusively dominates current listening pedagogy: passages of talk are listened to and examined for what has been stated. The similarly important prospective orientation is far less appreciated and almost entirely absent within ELT. Simply put, it involves listeners reacting to what is foreshadowed or *projected*. At a broad level, projection enables listeners to draw on schematic knowledge about likely trajectories that the conversation may take. This allows listeners to steer the direction of talk, formulate timely responses, take turns in multiparty talk, identify what action is underway, identify expected responses, and steel themselves for unwelcome news (Deppermann & Günthner ibid).

If this sounds somewhat familiar, it is important not to conflate projection with the conventional ELT concept of *prediction*. Prediction is widely conceived as a *learner strategy*, supported by various teaching practices (sometimes labelled *advance*

organizers), which focus on the identification of likely vocabulary, topic/theme and content within a forthcoming text (e.g. Graham & Santos 2015). Such predictions are fundamentally envisaged as aids to comprehension. They are enacted prior to listening, often with subsequent monitoring of assumptions as the text progresses. A variation is that of *anticipation*, which Field identifies as occurring '*when pauses in the discourse allow*' (Field 2008: 263, italics in the original).

In contrast, projection is neither a strategy nor done in the service of comprehension. Rather, it *is* an aspect of comprehension. The concept can be illustrated in a range of phenomena. Consider for example a situation in which a relative says 'What are you doing on Thursday night?' A competent listener recognizes this not as simply a request for information, but as almost certainly being preliminary to some further action, likely either a request (*Could you look after the kids?*) or an invitation (*Do you want to come to dinner?*). In sensing this, the listener is able to steer the direction of the conversation, most often by helping the action proceed (*Not much, why?*) or blocking it by formulating an account of being busy. Such projections have been attested in an enormous and varied range of empirical data within the conversation analysis literature.

Projection works through normative expectations of how language and interaction unfold given available cues. For instance:
- Intonation patterns have a discernible trajectory that allow us to estimate their end points
- Grammatical structures need certain elements to be considered complete (e.g. '*He put*' requires an object and an adverbial of place)
- Certain actions generate expectations of certain responses (e.g. offers being followed by acceptance or decline)
- *Dispreferred* responses (e.g. declining an invitation) are typically prefaced by elements such as pauses and *Oh*
- Multi-unit turns (e.g. telling a joke) need certain elements to be recognizably complete (e.g. a punchline)
- Event types may be characterized by certain regularities in the sequence of actions (e.g. how telephone calls begin and end; how service encounters unfold)

(Wong & Waring 2021)

The implications of projection are only beginning to be explored and have had very little impact on ELT to date. However, one implication is that because projected actions may be abandoned, modified, or thwarted, the relevance of 'being right' (in terms of anticipating what is actually on the recording) should be de-emphasized in favor of identifying a limited set of plausible trajectories.

> **Implication:** L-in-I activities should include a focus on what is being foreshadowed.

Participatory listenership

It has long been recognized that listeners often act upon the information they hear. This is often operationalized in classroom activities through students selecting items, making decisions, taking notes and so on. However, beyond this, Goffman's (1981) framework

highlights the more broadly participatory conduct required of listeners, through which they shape speaker behavior. These behaviors too are governed by normative expectations and may be noticeable in their absence. Among other things, listeners are normally expected to display attentiveness towards the speaker, signal their understanding, and are often expected to display a stance towards what is said (Stivers 2008). Key resources include:

- Various types of *minimal response* (e.g. *Mm hm*; *Okay*; *Yeah*)
- Assessments (e.g. *Wonderful*; *That's shocking*)
- Repair (*Huh?; Which one?*)
- Listener completions of the speaker's turn
- Nods, gestures, gaze, body position and so on

(Gardner 2001: 2)

Although most of these involve vocalizing, they function as displays of listener understanding and interpretation and can be rightfully addressed in listening classes. Any impression that they may be 'simple' is quickly disavowed by considering the first item, minimal responses, which Gardner (2001: 1) describes as 'exquisitely complex', a thesis more than justified in his nearly 300 pages dedicated mostly to examining eight uses of *Mm*.

Amongst the wider class of response tokens, Gardner (2001, chapter 2) distinguishes among *continuers* (e.g. *Mm hm*), *acknowledgment* tokens (e.g. *Yeah*), *news-marking* (e.g. *Oh*), and *change-of-activity* uses (e.g. *Okay*). Within each of these categories are multiple options that work in slightly different ways. Only some minimal responses will 'fit' as a plausible and non-problematic next turn and the wrong one may even cause offense. Different options can also steer the talk in different directions.

The case for incorporating response tokens into ELT listening has been well-established elsewhere, most cogently by Gardner (1998), who equated their value to that of the article and preposition systems. Further discussion here will therefore be left aside, except to note that these tokens vary cross-linguistically in form and use and are likely to require explicit teaching (Gardner 1988).

Beyond minimal responses, listeners are also expected to produce fuller turns that demonstrate their understanding and stance towards what they hear. Wong (2021) illustrates the consequences of problems in this domain through analysis of a phone call involving an L2 hearer, Huang and her friend Sally, who announces her need for cancer treatment. Although Huang claims understanding, her response to the announcement (*oh really*) and subsequent minimal turns are noticeably misaligned to the situation and unresponsive to Sally's evident need for sympathy and support. As the call comes to a close,

> 'Sadly, Huang does not return the well wishes, nor does she ask Sally to keep her apprised of the health crisis, nor does she express any desire to contact Sally to inquire about the radiation treatment and its side effects.' (Wong ibid.: 49)

Such listener responses, surely not intended in the way they are received, risk severely damaging crucial friendships. As this episode powerfully demonstrates, understanding

the content of talk is far from sufficient; effective hearers must also interpret what the speaker is doing and expecting, and then respond in ways that appropriately demonstrate their own stance.

> **Implication:** L-in-I activities should include a focus on responding as a listener, with particular attention to minimal responses.

Authenticity and recipient design

The positioning of students as addressees has a further implication for the design of listening texts. There are numerous well-reasoned and convincing calls for using 'authentic recordings wherever possible', with authenticity often defined as naturally occurring speech produced by and for proficient speakers rather than scripted specifically for teaching (Field 2008: 23). One such argument points to the very frequent appearance in published learning materials of overly careful enunciation, and consequent failure to represent the variable articulations found in natural speech. A second argument points to the numerous ways that the recurring elements of conversation are typically misrepresented in the scripting and performance of dialogues (Wong & Waring 2021).

Neither argument, of course, is tied to definitions of authenticity that specify 'by and for proficient speakers' (or indeed the now discredited 'by and for native speakers'), and it is this element of many definitions that most warrants re-evaluation. The principle of *recipient design* (Sacks & Schegloff 2007) holds that utterances are formulated primarily with the addressee in mind. For example, the same piece of advice might be expressed rather differently when talking to a child, a workmate or a stranger; similarly, and as long recognized, speakers modify their speech in ways that take account of assumptions about the language level of the L2 addressee, particularly in terms of speed, clarity of articulation, lexical choice and to some extent syntax. Thus, when learners are positioned as addressees, a recording may arguably be *more* authentic if the speech rate is reduced and the language simplified.

This observation provides a pragmatic and socially grounded basis for grading language in L-in-I. In short, recordings of dialogues could be relatively slow-paced, have apparently simplified language and yet be genuinely authentic and matched to the skill level of the students.

> **Implication:** where students are addressees, appropriate L-in-I recordings will include those graded to an approximate level of presumed competence.

Conclusion

For many language learners, particularly those in ESL contexts, everyday interactions are the most crucial site of listening, and their importance often far outweighs that of uni-directional texts such as television, podcasts and presentations. Successful L-in-I is underpinned by a number of complex skills and behaviours that are not widely appreciated by teachers and materials writers and it seems likely that when learners

bemoan the difficulty of listening outside the classroom, such factors represent a missing piece of the puzzle. The five principles of L-in-I outlined in this paper may be summarized as follows:

> Where possible, L-in-I texts and/or activities should:
> 1 Position learners as addressees or auditors
> 2 Include a focus on meaning as action
> 3 Include a prospective orientation
> 4 Involve participatory listenership
> 5 Be responsive to the audience

This list is far from definitive and indeed other principles have been touched on but not pursued in this paper, including the role of multimodalities and orienting away from the expectation of a single 'right answer'. At first glance, it may seem somewhat daunting to reimagine listening pedagogy in ways that might accommodate these, and some teachers may baulk at calls to transform their practice or be otherwise unconvinced of the practicalities. Others, however, will recognize such concerns as simply representing technical and procedural challenges and will recall ELT's long traditions of innovation and creativity in adapting to change.

References

Deppermann, A. and **Günthner, S.** 2015. 'Introduction: Temporality in Interaction' in A. Deppermann & S. Günthner (eds.). *Temporality in Interaction, pp. 1-23*. Amsterdam: John Benjamins.

Enfield, N. J. and **Sidnell, J.** 2017. *The Concept of Action*. Cambridge: Cambridge University Press.

Field, J. 2008. *Listening in the Language Classroom*. Cambridge: Cambridge University Press.

Gardner, R. 1998. 'Between Speaking and Listening: The Vocalisation of Understandings'. *Applied Linguistics, 19*/2: 204-224.

Gardner, R. 2001. *When Listeners Talk: Response Tokens and Listener Stance*. Amsterdam: John Benjamins.

Goffman, E. 1981. *Forms of Talk*. Philadelphia, PA: University of Pennsylvania Press.

Graham, S. and **Santos, D.** 2015. *Strategies for Second Language Listening: Current Scenarios and Improved Pedagogy*. Basingstoke: Palgrave Macmillan.

Nation, I. S. P. and **Newton, J.** 2009. *Teaching ESL/EFL Listening and Speaking*. New York: Routledge.

Sacks, H. and **Schegloff, E. A.** 2007. 'Two preferences in the organization of reference to persons in conversation and their interaction' in N. J. Enfield & T. Stivers (eds.). *Person Reference in Interaction: Linguistic, Cultural, and Social Perspectives*, *pp.* 23-28. Cambridge: Cambridge University Press.

Siegel, J. 2015. *Exploring Listening Strategy Instruction through Action Research*. Basingstoke: Palgrave Macmillan.

Stivers, T. 2008. 'Stance, Alignment, and Affiliation During Storytelling: When Nodding is a Token of Affiliation'. *Research on Language and Social Interaction, 41/1:* 31-57.

Vandergrift, L. 1997. 'The Cinderella of Communication Strategies: Reception Strategies in Interactive Listening'. *Modern Language Journal*, *81/4:* 494-505.

Vandergrift, L. 2007. 'Recent Developments in Second and Foreign Language Listening Comprehension Research'. *Language Teaching, 40*/3: 191-210.

Wong, J. 2021. 'Our Storied Lives: Doing And Finding Friendship II' in J. Wong & H. Z. Waring (eds.). *Storytelling in Multilingual Interaction: A Conversation Analysis Perspective, pp. 41-54*. New York: Routledge.

Wong, J. and **Waring, H. Z.** 2021. *Conversation Analysis and Second Language Pedagogy: A Guide for ESL/EFL Teachers* (2nd ed.). New York: Routledge.

The author

Jonathon Ryan is a language teacher and researcher based at Wintec in Hamilton, New Zealand. His research focuses mostly on second language pragmatics and on applying insights from conversation analysis to developing teaching materials.

Email: jonathon.ryan@wintec.ac.nz