English Language and Linguistics, 1–32.
doi:10.1017/S1360674324000625

Noticing, identifying and discriminating sociolinguistic variants in England[1]

ERIK SCHLEEF, EVELYN N. ROTH, BRADLEY MACKAY and JANA PFLAEGING
University of Salzburg
(Received 14 November 2023; accepted 28 October 2024)

This article explores the extent to which listeners vary in their ability to notice, identify and discriminate variable linguistic features. With a view to improving speaker evaluation studies (SES), three types of experiments were conducted (*noticing tasks*, *identi cation tasks* and *discrimination tasks*) with regard to variable features using word- or sentence-based stimuli and focusing on three variables and their variants (ING): [ŋ], [n]; (T)-deletion: [t], deleted-[t]; (K)-lenition: [k], [x]. Our results suggest that the accurate *noticing*, *identifying* and *discriminating* of variants is somewhat higher in words than in sentences. Correctness rates di er drastically between variants of a variable. For (ING), the non-standard variant [n] is more frequently identified and noticed correctly. Yet, for the variables (T)-deletion and (K)-lenition, the standard variants are identified and noticed more successfully. Results of the current study suggest that a more rigorous elicitation of *identi cation* and *noticing* abilities might be useful for a more complete understanding of the nature of social evaluation.

Keywords: noticing, discrimination, identification, speaker evaluation, England

1 Introduction

The last two decades have seen a steep increase of speaker evaluation studies (SES) that investigate how listeners evaluate speaker guises which di er in single variable features, for example, (ING) with its variants [ŋ] and [n]. Many studies have shown that speakers are downgraded socially when they use forms that are more marked in a specific context (e.g. Campbell-Kibler 2007; Labov *et al.* 2011), such as

[1] We would like to thank two anonymous reviewers for their insightful comments on an earlier version of this manuscript. We also gratefully acknowledge the funding that our study received from the European Union s Horizon 2020 research and innovation programme under the Marie Sk odowska-Curie grant agreement no. 754340.

non-standard [n] in newscaster speech. A large variety of factors has been shown to a ect evaluation, including speaker and listener socio-demographic factors (Wagner & Hesson 2014; Schleef & Flynn 2015), specific feature combinations (Pharao *et al.* 2014; Schleef 2023), region (Schleef *et al.* 2017), the social and linguistic context in which a form occurs (Bender 2005; Vaughn 2022a), cognitive style (Wagner & Hesson 2014; Levon & Buchstaller 2015), topic (e.g. Campbell-Kibler 2009), etc.

This article focuses on a factor that has so far not been explored in much detail: the extent to which listeners vary in their ability to *notice*, *identify* and *discriminate* variable linguistic features. With the help of these concepts, we seek to get an analytical grip on several related phenomena of speech perception. The ability to *notice* variable features refers to a listener s ability to become aware of a given variant without any previous indication of it. The ability to *identify* and *discriminate* variants, in turn, describes participants capacity to hear a given variant correctly when their attention has been purposely drawn to a focus word prior to listening to a stimulus.

We do not mean to suggest that listeners *need* to be able to *notice*, *identify* and *discriminate* variation to evaluate it as part of SES (see, e.g., Campbell-Kibler 2011; De 2017). Rather, we hypothesise that individual listeners ability to do this may be variable, which may have *some* e ect on evaluation. We test what the extent of this variability is to determine its (potential) impact on the response patterns observed in SES. If we find evidence for low *noticing*, *identi cation* and *discrimination* rates, researchers may want to consider collecting such information and include it as a factor in statistical models, as null results in SES may otherwise conflate two separate phenomena: variation that listeners simply do not notice, and variation that listeners notice, but have no clear evaluation of.

To test listeners ability to *notice*, *identify* and *discriminate* variable pronunciation, we focus on three features which are (i) variable in England and (ii) di er somewhat in their sociolinguistic status: (ING), (T)-deletion and (K)-lenition. To account for the conceptual di erences between the ability to *notice*, *identify* and *discriminate* variants, we implemented several listening tasks and collected responses from 160 listeners in England. We thereby seek to determine the proportion of correct response rates per task type, variable, variant and carrier type (word/sentence), and explore factors that a ect participants responses. More generally, we aim to raise awareness of listeners ability to *notice*, *identify* and *discriminate* variable linguistic features as a dimension of SES that potentially has statistical relevance to listeners evaluation patterns.

Before we present the results of our study, we provide some background on SES and the variables. We introduce the concepts of *noticing*, *identifying* and *discriminating*, and explain how these concepts relate to our task designs. We then present our research questions and hypotheses regarding participants response patterns.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

2 Background

2.1 Speaker evaluation studies (SES)

In SES, participants listen to several sets of recordings from either di erent speakers (verbal-guise technique, Cooper 1975) or the same speaker (matched-guise technique, Lambert *et al.* 1960), which they are then asked to evaluate. Conducting an SES normally involves recording speakers and then manipulating the recordings, for example, by cutting one variant out and pasting another variant into the guise. The recordings di er in the aspects that the researcher is interested in, such as dialects (London versus Liverpool English) or specific features (e.g. non-standard versus standard features, such as [n] vs [ŋ] in unstressed syllables).

Survey participants are then asked to rate each individual guise on, for instance, the perceived accuracy, educatedness or intelligence of the speaker. SES normally use scales to record the assessment, such as semantic di erential scales with bipolar adjectives (Osgood *et al.* 1957). The fact that participants are asked to evaluate the speaker rather than the speaker s speech is key; this way of collecting evaluations is supposed to access language attitudes in a more disguised manner by not focusing on the language but the speakers of certain varieties.

SES at variant level have been used quite widely in the last few years as they allow us to answer questions that have emerged in recent strands of sociolinguistics, for example, regarding the interplay of psychological and sociolinguistic factors and third-wave variationist sociolinguistics (Eckert 2012). In particular, variation in (ING) and (T)-deletion have been an important focus of these studies, and we now provide more detail on these variables.

2.2 The linguistic variables

2.2.1 Variation in (ING)

In most English dialects, (ING) is stable and has minimally two realisations in unstressed positions: [ŋ], e.g. *talking*, and [n], e.g. *talkin* (Wells 1982: 262 3; Labov 2001; Hazen 2006; Tagliamonte 2004). Both the sociolinguistic constraints operating on (ING) and the social meanings attributed to these variants are relatively well described (e.g. Campbell-Kibler 2007, 2009, 2011; Labov *et al.* 2011; Schleef & Flynn 2015; Schleef *et al.* 2017; Bailey 2019). For example, Campbell-Kibler (2011) shows that [ŋ] guises are normally rated as intelligent, educated, articulate and less likely to be a student, while [n] guises are considered informal and less likely to be gay-sounding.

Previous research has shown variation in (ING) to be constrained by both internal and external factors (e.g. Trudgill 1974; Labov 2001; Hazen 2006; Schleef *et al.* 2011). Among the internal constraints are phonological conditioning e ects, for example, [ŋ] is more likely to occur when the following segment begins with a velar sound (Houston 1985: 19 20). In previous segments, a velar stop or apical show

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

a strong tendency to prompt the apical variant [n]. Also, stress placement predicts the production of [ŋ] or [n] (Hazen 2006: 583) as well as morphological class. Verbs tend to favour [n], whereas [ŋ] typically occurs with nouns (Houston 1985; Labov 2001: 86 7; Tagliamonte 2004; Vaughn 2022a: 513).

External constraints on (ING) include style, gender and socioeconomic class. The variant [n] is normally more frequent in informal style. It is often used more frequently by men, and generally, there is more use of [n] as one goes down the social scale; however, all of these constraints interact with region and the baseline frequency of [n] in a particular area (e.g. Labov 2001: 90). Based on VARBRUL probabilities, Houston (1985: 108) suggests there are two dialect groups in Britain regarding (ING): the southern or internal group, favouring [ŋ], and the northern and peripheral group, favouring [n]. Houston (1985: 103), focusing on working-class (WC) speech, cites [ŋ] values of 20 to 42 per cent for London, 21 per cent for Manchester and 18 per cent for Edinburgh. Tagliamonte (2004) found little symbolic social value attached to (ING) in York, England. Social class is the only social factor that reaches statistical significance, and its e ect size is not particularly strong. This is very di erent from the more extreme sociolinguistic conditioning in the US, Australia and, notably, southern England (Tagliamonte 2004: 401).

2.2.2 (T)-deletion

In (T,D)-deletion, the English morpheme-final post-consonantal plosives /t/ or /d/ are deleted when followed by a consonant, vowel or pause. Similar to (ING), it occurs in almost all varieties of English. Much less is known about it in British English (Tagliamonte & Temple 2005; Pavlík 2017; Baranowski & Turton 2020) than in American English (e.g. Wolfram 1969; Fasold 1972; Guy 1980, 1991; Labov 1989; Santa Ana 1996; Bayley 1994; Coetzee & Kawahara 2013). (T,D)-deletion is constrained most strongly by the phonological context and the morphological status of the variable. Following consonants and non-sonorous segments favour deletion; vowels and sonorants inhibit it (Labov 1989; Guy 1991; Tagliamonte & Temple 2005; Smith *et al.* 2009; Hazen 2011). Preceding sibilants favour deleted-[t], followed by stops, nasals, other fricatives and liquids (Labov 1989). In addition, (T)-deletion is predicted by higher-level linguistic constraints, such as word class and the morphological structure of a word (Guy 1980, 1991; Tagliamonte & Temple 2005; Baranowski & Turton 2020; Mackay 2024). In particular, underived, mono-morphemic words have been shown to favour deleted-[t] (Tagliamonte & Temple 2005: 284, 290), whereas inflected verb forms disfavour it (Guy 1991; also Pavlík 2017: 197).

Previous research has focused only rarely on (T)-deletion in England. Tagliamonte & Temple (2005: 287 8) found (T,D)-deletion to occur in 24 per cent of their York data. Baranowski & Turton (2020: 15) report rates of 37 per cent for males and females in Manchester. Pavlík gives a total of 26 per cent for words ending in *-n t*, *-nt* and *-ed* in the Standard British English of sixteen BBC-newscasters. Social factors do matter, but

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

this varies from study to study. In York, males delete slightly more frequently (Tagliamonte & Temple 2005). There is no gender di erence in Manchester (Baranowski & Turton 2020: 15 16). Social class was also not significant in Manchester, but there are stylistic di erences with less deletion in more formal styles.

2.2.3 (K)-lenition

(K)-lenition is regionally restricted and stereotypically associated with Liverpool English. It occurs word-finally and word-medially after vowels. When lenited after low and back vowels, (K) is realised as an a ricate [kx] or a fricative [x], as in *back* [bax] (Watson 2006: 58). Palatal fricatives occur after [i , ei, ai] (Watson 2007: 353). Although no precise and current usage rates are available, its status as a non-standard feature likely means that it is used further down the socioeconomic scale.

2.3 Variable linguistic features, perception and salience

2.3.1 Language-external and -internal explanations of salience

The notion of *salience* intuitively relates to phenomena of prominence, awareness or surprise, but it has often been found to be underspecified in academic discourse, including linguistics (Boswijk & Coler 2020: 713; also Zarcone *et al.* 2016). More recently, scholars have sought to synthesise conceptualisations of salience in di erent fields of linguistics, such as sociolinguistics and cognitive linguistics (e.g. Boswijk & Coler 2020), to arrive at a working definition that accounts for several important dimensions of salience.

Among these dimensions are *language-internal* and *language-external* aspects that can explain why certain linguistic features stand out to certain (groups of) language users. Language-internal explanations include the 'presence of phonological contrast, great phonetic distance, internally defined naturalness, semantic transparency, or a particular syntactic or prosodic environment (Kerswill & Williams 2002: 105, also quoted in Boswijk & Coler 2020). Language-external explanations, in turn, comprise 'cognitive, pragmatic, interactional, social psychological and socio-demographic factors (Kerswill & Williams 2002: 105, also quoted in Boswijk & Coler 2020).

In SES, empirical work has focused on language-external factors and the social meanings of variants (Levon & Fox 2014; also Labov 2001: 25 8). While language-internal factors are equally at work (see Trudgill 1986: 11, also quoted in Boswijk & Coler 2020: 716), their exact contribution to a feature s salience has not been sufficiently tested.

Our study shifts the focus more clearly to the question of whether or not participants can reliably *notice*, *identify* and *discriminate* di erent variants, as reflected in the task types implemented. Slight di erences in task design serve to explore various degrees of salience and the potential e ect of drawing respondents attention to a variable.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

2.3.2 *Noticing, identifying and discriminating variants*

Noticing. While terms may vary, the *noticing* of particular language use and connecting it to social meanings is common to various models of speech perception (e.g. Kristiansen 2008; Purschke 2011). The understanding of *noticing* adopted here may be best explained by referring to Preston s work on *language regard.*

Preston (2017) has developed a processual model of where and why variability may emerge in the social evaluation process.[2] In a crucial first step, listeners must *notice* an instance of language use; otherwise, they cannot react to it. A variant may be noticed because it di ers from the listener s own preferred or expected realisation of the item. The noticed language instance is subsequently 'classified according to social, contextual or linguistic criteria, and then 'imbued with evaluative information drawn from the listener s stored cognitive representations of the classification. In a final step, there is a deliberative or automatic response.

The model s exact workings will not be tested here. Instead, we focus on the first step of *noticing* and examine to what extent individuals *notice* the di erent variants of a variable. In particular, our *noticing tasks* were designed in such a way that, in the question prompt, respondents attention is *not* initially drawn to a particular focus word or feature; instead, they first listen to a recording and are only afterwards asked to report which pronunciation variant was heard.

When it comes to interpreting the results, we believe that responses to our *noticing tasks* reflect the complexity of natural speech perception, with both language-internal and -external factors influencing participants ability to *notice* a variant. As argued above, Preston s (2010) conceptualisation of *noticing* provides important insights here. Crucially, he argues that *noticing* is a dynamic process that depends on situations, tasks and properties of individuals, which acknowledges the role of elicitation conditions and listeners procedural capacities, such as working memory (Bassili & Brown 2005: 553 4).

Di erent theoretical positions are evident in the literature as to how conscious the process of *noticing* is. According to Nycz, *noticing* (in second language learning research) refers to 'conscious awareness and subjective experience of a linguistic feature; it is a step beyond mere *perception* of a feature (Nycz 2016: 64, quoted in Montgomery & Moore 2018, original emphasis). Preston (2017: 20), on the other hand, argues that the *noticing* of a variant may be conscious or unconscious. We subscribe to this view and note that our task design does not support any conclusions as to how consciously or unconsciously a listener *noticed* a given variant. We also believe that evaluation is possible with or without a feature being *noticed* consciously and with or without a listener completing a set of *noticing tasks* successfully (see Campbell-Kibler 2011; De 2017). However, the ability to *notice* a feature and indicate the correct variant *may* influence evaluation; we hypothesise that

[2] A similar related model of how linguistic variants acquire social meaning has been proposed by Silverstein (2003). As our focus here is predominantly on questions of *noticing*, *identifying* and *discriminating* rather than aspects of meaning ascription and evaluation, we refrain from discussing this model further.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

evaluation may be stronger among those who perform better in *noticing tasks* than those who struggle with them.

Identifying. The notion of *identifying* a feature refers to a participant s ability to indicate precisely which variant they heard (e.g. answer options: *singing* or *singin*) in a recording of a word or sentence (e.g. Drager 2018: 75 8; Grover *et al.* 2021). Typically, in so-called *identi cation tasks*, listeners are told they will hear words or sentences, sometimes their attention is drawn to a particular focus word, and they are subsequently asked to indicate what variant they heard (e.g. represented as a written word). In some cases, target words or sounds are purposely kept ambiguous (e.g. Squires 2014). Our implementation followed what is commonly done, that is, drawing listeners attention to a focus word before hearing a stimulus, and presenting them with the orthographic equivalents of the respective variants from the start.

Identi cation tasks can be used for various research questions, for instance, to investigate the e ect of specific social primes, or to test intelligibility (Hay *et al.* 2006; Drager 2018: 75 8). Given that many SES presuppose participants ability to correctly *identify* variants, we used *identi cation tasks* to test how accurately participants can *identify* di erent variants (e.g. [ŋ] or [n]) of specifically prompted parts of speech. The ability to hear a variant correctly and assign the correct answer option *may* influence listener evaluation; evaluation *may* not be possible if a listener cannot correctly *identify* a variant. Alternatively, evaluation may occur but it may be weaker when compared to listeners with very high *identi cation* results.

Discriminating. By *discriminating*, we mean a participant s ability to distinguish between two variants of a variable, although only *indirectly* by indicating whether a certain type of sound in two di erent recordings was pronounced the *same* or *differently.* This understanding builds on previous work, where speech perception data was collected by means of *discrimination tasks* (e.g. Drager 2018: 78 80). Our exact experimental set-up di ered slightly from other *discrimination tasks* commonly used to examine mergers in progress (Drager 2018: 78 80), in as far as we pointed out the focus word in writing before participants heard the recordings and were asked to classify them as either *same* or *different* (or *I don t know*). This was done deliberately, however, as we sought to test whether participants had the ability to *discriminate* between variants in an unambiguous setting without any fillers or added confounding variables.

Given this experimental design, *discrimination* and *identi cation tasks* showed many similarities. However, they also di ered in a few key aspects: in the *identi cation tasks*, participants only heard one realisation of the variable in question, while in the *discrimination tasks* they heard two (e.g. [ŋ]/[n], [ŋ]/[ŋ] or [n]/[n]) right after each other, giving direct comparability. Also, *discrimination tasks* seemed to demand a lower cognitive load, as the task only requires participants to compare and evaluate acoustic input regarding the sameness of sounds, but not to additionally match it with the corresponding written representation of a given variant.

Table 1 provides a comparative summary of what precisely the speech-perception activities of *noticing*, *identifying* and *discriminating* involve. In particular, we make

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 1. *Overview describing the nature of the speech-perception activities of* noticing, identifying *and* discriminating *linguistic variants*

	Noticing	Identifying	Discriminating
Attention drawn to variants of a variable	lowest degree: • none to low attention drawn to variant in actual task • some attention drawn to variants in survey instructions	highest degree: • attention drawn to variant by question text and visible answer options • some attention drawn to variant in survey instructions	some degree: • attention drawn to standard variant by question text, but not by answer options (‘same , ‘di erent , ‘I don t know) • some attention drawn to variant in survey instructions
Salience	impacted by language-external and -internal factors	more impacted by language-internal factors	more impacted by language-internal factors
Assumed cognitive load	sentences: high • barely any attention drawn to variants increases load	words: lower than in sentence-based *identi cation tasks* sentences: higher than in word-based *identi cation tasks* • highest degree of attention drawn to variants lowers load • listeners need to match *visual* representation with *acoustic* input increases load	words: lower than in sentence-based *discrimination tasks* sentences: higher than in word-based *discrimination tasks* • some attention drawn to variants lowers load • listeners do not need to match *acoustic* input with *visual* representation of variant lowers load • first realisation has to be held in working memory and then compared to second realisation increases load
Naturalness	most natural • full sentences • no knowledge of focus word prior to listening barely any attention drawn to variants	less natural than *noticing tasks*	less natural than *noticing tasks*
Assumed level of difficulty	most challenging	semi-challenging	least challenging

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

assumptions about the ways in which the di erent task types draw participants attention to variants, what types of salience are raised, and we hypothesise about the cognitive load the tasks demand. We specifically acknowledge that, in the present study, participants were introduced to variants at the beginning of the survey or in the case of the *identi cation* and *discrimination tasks* were explicitly told about focus words and sounds, which leads to a certain degree of *introspective awareness* on the part of the listener. Raising such introspective awareness happens deliberately in *identi cation* and *discrimination tasks*, so that responses mainly reflect the ability to listen selectively for certain variants, classify them or evaluate their sameness. In *noticing tasks*, however, we sought to avoid an unnaturally high level of introspective awareness to collect perception data that is most complementary to data typically elicited through SES. Taking such assumptions into account, the table also indicates the assumed level of naturalness and difficulty for respondents. Such a comparative account serves to pinpoint similarities and di erences between task types and to ultimately be able to draw more nuanced conclusions about what precisely each of them tests, how performance di erences ought to be interpreted and which task type lends itself best to an implementation in SES.

2.4 Research questions and hypotheses

This study seeks to test respondents ability to *notice*, *identify* and *discriminate* variable linguistic features at word and sentence level. We focus on three features which are (i) variable in England and (ii) di er somewhat in their sociolinguistic status:

- variation in (ING) e.g. *talking* realised with [ŋ] or [n],
- variation in (T)-deletion e.g. *went* realised with [t] or deleted-[t],
- variation in (K)-lenition e.g. *padlock* realised with [k] or [x].

We ask the following questions:

- What is the proportion of correct *noticing*, *identifying* and *discriminating* of variants of (ING), (T)-deletion and (K)-lenition at word and sentence level?
- What task-dependent factors and listener characteristics a ect correct *noticing*, *identifying* and *discriminating*?
- Would it be worthwhile to include tasks that assess respondents ability to *notice*, *identify* and *discriminate* variable linguistic features in SES to then use this information as a factor in the statistical analyses? If so, which task type is most suitable?

We hypothesise that the highest proportion of correct answers will be achieved for the *discrimination* of variants, followed by the *identi cation* and finally the *noticing* of variants. This is because the tasks di er in cognitive demand (see table 1). In addition,

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

we expect more correct answers when a variant occurs in a single word rather than in a sentence, due to the additional cognitive load associated with processing sentences. Considering that the nature of the tasks is relatively independent of gender and social class, we do not expect these listener characteristics to influence the number of correct responses. Likewise, we do not expect region to a ect respondents correctness rates significantly since we collected data from several regions in the North of England (not only Liverpool), so that familiarity with the [x]-variant was not a given. Additionally, all participants were briefly introduced to all relevant variants at the beginning of the survey, which may decrease the e ect of region further.

3 Methods

3.1 Guise creation

In a first step, we created several sets of stimuli words. To ensure comparability and to exclude confounding factors, focus words and contexts follow several criteria: all target words consist of two syllables with the feature of interest occurring in the final coda. All word-final (T)-deletion targets are preceded by /n/; all word-final (K)-lenition features are preceded by back vowels and the non-standard form was recorded as [x]. In a next step, words were integrated into sentences, occurring as the third or fourth word of sentences that consist of 10 14 syllables in total. In sentences, features of interest are always followed by a vowel (see the Appendix for all text materials).

Words and sentences were recorded in a sound-attenuated booth at the University of Salzburg, with one female speaker in her mid-40s originally from Liverpool. Her accent is clearly Northern but may not allow listeners to place her in a specific city. We only worked with unmanipulated recordings of words and sentences since (i) we did not conduct an SES and therefore did not need to hide which features we were interested in, and (ii) recent research has questioned whether the cut-and-paste technique is suitable for all experimental goals (Vaughn 2022b).

3.2 The survey

In total, we had five di erent tasks: *noticing* linguistic variants in sentences (2 blocks), *identifying* variants in words, *identifying* variants in sentences, *discriminating* variants in words and *discriminating* variants in sentences. To circumvent any fatigue e ects, we split these tasks into two surveys (table 2).

To counteract possible confirmation bias or lexical e ects, each survey came in two versions, which only di ered in what variant participants would hear with each individual word or sentence. The surveys were created in *LimeSurvey* (LimeSurvey GmbH 2023). An information page ensured that only participants with working audio equipment and without any hearing impairments took part in the survey. A subsequent training section explained what the survey entailed and that it was about ‘identifying

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 2. *Overview of surveys*

Survey 1	Survey 2
noticing t sks (sentences) 1st block	***noticing t sks*** (sentences) 2nd block
identi c tion t sks (words)	***identi c tion t sks*** (sentences)
discrimin tion t sks (sentences)	***discrimin tion t sks*** (words)

sounds. Each of the three variables was briefly introduced here to familiarise participants with the written representation of the non-standard forms: for the (ING)-variants, we used *-ing* and *-in* , respectively. The variants of (T)-deletion were represented orthographically as *-t* and - (as in *presen*), respectively. Finally, the spelling variants for (K)-lenition were *-ck* and -[x], respectively.[3] Participants also had the opportunity to listen to audio examples of the variants. This feature introduction is a potential weak point of the study, as we had to raise some awareness of the variables and variants. This is unproblematic for the *discrimination* and *identi cation tasks* but undesirable for the *noticing tasks*, as we did not want participants to be aware of the target feature before hearing the stimuli. To mitigate this risk, it was suggested to participants that they 'might be asked about a feature on the training page. And, indeed, the survey contained other features as well (e.g. variation in (W) and (V) in the trial questions) to further draw away attention and avoid training e ects.

The introductory part was followed by the instructions for the first block, a trial question and, eventually, the nine randomised sentences of the *noticing tasks*. The *noticing tasks* also included three fillers and an attention check. Next followed two further experimental blocks with *identi cation* and *discrimination tasks*, also with attention checks, but no fillers. At the beginning of each part, participants were reminded to always choose the answer that corresponds to what they actually heard rather than what they would expect based on Standard English. The order of test tokens was randomised throughout. The final section collected basic biographical and experimental information (see Appendix).

Proli c (Prolific 2023), a subject pool for online experiments, was used for data collection. We used its screening options to include 80 participants each, who were born and still live in the North of England (from the North East, North West or Yorkshire and the Humber) and the South (from London, South East or the South West). We recruited a total of 160 participants (50 per cent women, 50 per cent men), of whom 48.8 per cent self-identified as working class and 51.2 per cent as middle class.[4] We admitted participants between 18 and 100 years of age (mean age=44.25 years, sd=13.8 years, min=20, max=76) whose L1 is English and who are UK nationals who

[3] Representing (K)-lenition proved particularly difficult, and we researched strategies of presenting [x] used in popular media discourse. We considered and tested several options, including '< and 'ch , and concluded that [x] would serve our purposes best when combined with a training phase.

[4] Two respondents self-identified as lower upper class. To include this data in the statistical analysis in any meaningful way, we merged this category with 'middle class .

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 3. *Structure of a noticing task*

Step	Information visible on the screen
1	Recording
2	You have just heard a recording of the word *walking.* How was it pronounced? • A: walking • B: walkin • C: I don t know.

spent most of their time in the UK before turning 18. Participants could take part in only one of the surveys, which took 15 minutes on average, and they were paid £3.50 each. After survey completion, incomplete responses and participants who failed multiple attention checks were removed. Data files were then prepared for the statistical analysis in R (see section 3.6). Sections 3.3 to 3.5 provide more detail on the three experimental tasks.

3.3 Noticing tasks

We conducted sentence-based *noticing* experiments in two randomised blocks, each containing three tokens of each of the three variables. Participants first answered a trial question to familiarise themselves with the question format. They then clicked *Next* to listen to the first test sentence; once the recording had played through, the audio player disappeared automatically to ensure that the recording could only be listened to once (step 1, table 3). Participants would then be asked whether a specific word in this sentence was pronounced in one way or another, that is, the target word was pointed out to participants only after the stimulus was heard (step 2, table 3).[5] We only used sentences as stimuli as this method most closely resembles what happens in SES.

3.4 Identi cation tasks

The *identi cation tasks* were very similar to the *noticing tasks* except that the target word was pointed out to the participants *prior* to hearing the recording. We conducted word- and sentence-based *identi cation tasks* in separate blocks and surveys, each containing six tokens of each of the three variables. After a trial question, participants clicked *Next* and read that they would hear a recording of a focus word after clicking the play-button, and that they would be asked how the focus word had been pronounced, choosing one out of three answer options (table 4).

[5] It follows from this description that *noticing tasks* required an act of *indicating* which variant was heard (by selecting an answer option). While this aspect is perhaps more inherent to notions of *identifying* and *discriminating* rather than *noticing* (see section 2.3.2), it is something that could hardly be avoided with the type of experiment conducted here.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 4. *Structure of an identification task*

Step	Information visible on the screen
1	You will hear a recording of the word *present.* Recording How was it pronounced? • present • presen • I don t know.

Table 5. *Structure of a discrimination task*

Step	Information visible on the screen
1	Focus on the final sound in the word *notebook* in both recordings. Recording Click SAME if you think the final sound was the same both times. Click DIFFERENT if you think it sounded di erent. • SAME • DIFFERENT • I don t know.

3.5 Discrimination tasks

Finally, we implemented word- and sentence-based *discrimination tasks* in separate blocks, each containing six tokens of each of the three variables. After a trial question, participants clicked *Next* and read that they are asked to focus on a specific sound in a focus word contained in two separate recordings and to decide whether it sounded the same or di erent in both instances (table 5).

3.6 Statistical analysis

Correctness rates were submitted to mixed-e ects logistic regression models built with the *glmmTMB* package (Brooks *et al.* 2017) in R (R Core Team 2021). Separate models were built to investigate the three variables (i.e. (ING), (T)-deletion, (K)-lenition) and the three task-types (i.e. *noticing*, *identi cation* and *discrimination tasks*). Responses were dummy coded with 0 for incorrect answers and 1 for correct answers. Models were fitted with random e ects for ARTICI ANT and maximum models were manually stepped down to include factors which significantly improved the model fit. The following overview is an alphabetical list of the factors and levels considered when fitting the models:

- AGE: younger (35) | middle (36 65) | older (≥66)
- AUDIO: headphones | in-built computer speakers | external speakers

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

- CLASS: working class | middle class
- GENDER: man | woman
- REGION: North | South
- VARIANT: standard | non-standard

4 Results

4.1 Degree of correctness: noticing, identifying and discriminating

Approaching our first RQ, table 6 and figure 1 show the overall proportion of correct responses (per variable, variant and task type).[6] L1 speakers of English do not *discriminate* and *identify* all variants correctly and less surprisingly do not show a 100 per cent *noticing* rate. Also, correctness rates vary within variables, which is why our focus will be on the variants of the three variables. Overall, we observe higher correctness rates in the *discrimination tasks* (never below 80 per cent). Correctness rates for *noticing* and *identifying* vary widely. Given the variation between task type and variant, we will now explore the factors influencing variation in these data based on binomial logistic regressions (tables 7 15).

4.2 Factors influencing correct noticing, identifying and discriminating

Noticing tasks. Results (table 7) indicate that respondents are significantly better at *noticing* the non-standard variant [n] (90 per cent) compared to the standard variant [ŋ] (72 per cent) (est.=-1.42, $p< 0.001$). For both (T)-deletion and (K)-lenition, this pattern is reversed, with participants being able to correctly respond to *noticing tasks* at a higher rate when hearing the standard variants [t] (est.=1.15, $p<0.001$, table 8) and [k] (est.=0.97, $p<0.001$, table 9) as opposed to the non-standard variants (deleted-[t] and [x], respectively). When hearing non-standard [x], only 70.8 per cent of respondents answered correctly one of the lowest scores overall. While it did not prove a significant predictor of respondents CORRECTNESS rates, including REGION as a factor in the (T)-deletion model improved the overall model fit.

Identi cation tasks. As with the *noticing tasks*, participants performed significantly better when required to *identify* the non-standard variant [n] as opposed to the standard variant [ŋ] (est.=-0.60, $p=0.043$, table 10). Results also indicate a significant e ect for CARRIER. When the stimulus consisted of a word, respondents were significantly better at *identifying* the variant correctly than when they were presented with a carrier sentence (est.=1.10, $p=0.021$, table 10). Working-

[6] We are aware that CORRECTNESS rates are potentially confounded with bias. If respondents decided to always select the same variant, their CORRECTNESS rates would be much higher for that specific variant, which does not reflect their actual ability to *notice*, *identify* or *discriminate* variants. As there is no reason to assume that respondents would not answer questions to the best of their ability or that such a bias would be variant-specific, we believe that the CORRECTNESS rates we found are a relatively accurate representation of participants ability to *notice*, *identify* and *discriminate* a given variant.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 6. *Proportion of correct responses per* TASK TYPE, VARIABLE *and* VARIANT, *with number of correct responses (*noticing tasks *and* identification tasks*: out of a total of 240,* discrimination tasks*: out of a total of 480) as well as upper and lower 95 per cent confidence intervals in brackets*

		(ING)		(T)-deletion		(K)-lenition	
	VARIANT / CARRIER	[ŋ]	[n]	[t]	deleted-[t]	[k]	[x]
Noticing	sentence	**72** (n=172) (65.6 77)	**90.0** (n=216) (85.5 93.2)	**94.2** (n=226) (90.4 96.5)	**83.8** (n=201) (78.5 87.9)	**85.8** (n=206) (80.8 89.7)	**70.8** (n=170) (67.7 76.2)
Identifying	word	**79.2** (n=190) (73.5 83.9)	**95.8** (n=230) (92.4 97.8)	**100** (n=240) (0)	**93.3** (n=224) (89.4 95.9)	**97.9** (n=235) (95.1 99.1)	**73.3** (n=176) (66.3 78.6)
	sentence	**83.8** (n=201) (78.5 87.9)	**89.6** (n=215) (85 92.9)	**97.5** (n=234) (94.5 98.9)	**64.2** (n=154) (57.9 70)	**95.4** (n=229) (91.9 97.5)	**79.2** (n=190) (73.5 83.9)
Discriminating	word	**94.6** (n=454) (92.2 96.3)		**80.8** (n=388) (77 84.1)		**88.5** (n=425) (85.4 91.1)	
	sentence	**85.6** (n=411) (82.2 88.5)		**84.0** (n=403) (80.4 87)		**91.2** (n=438) (88.4 94.5)	

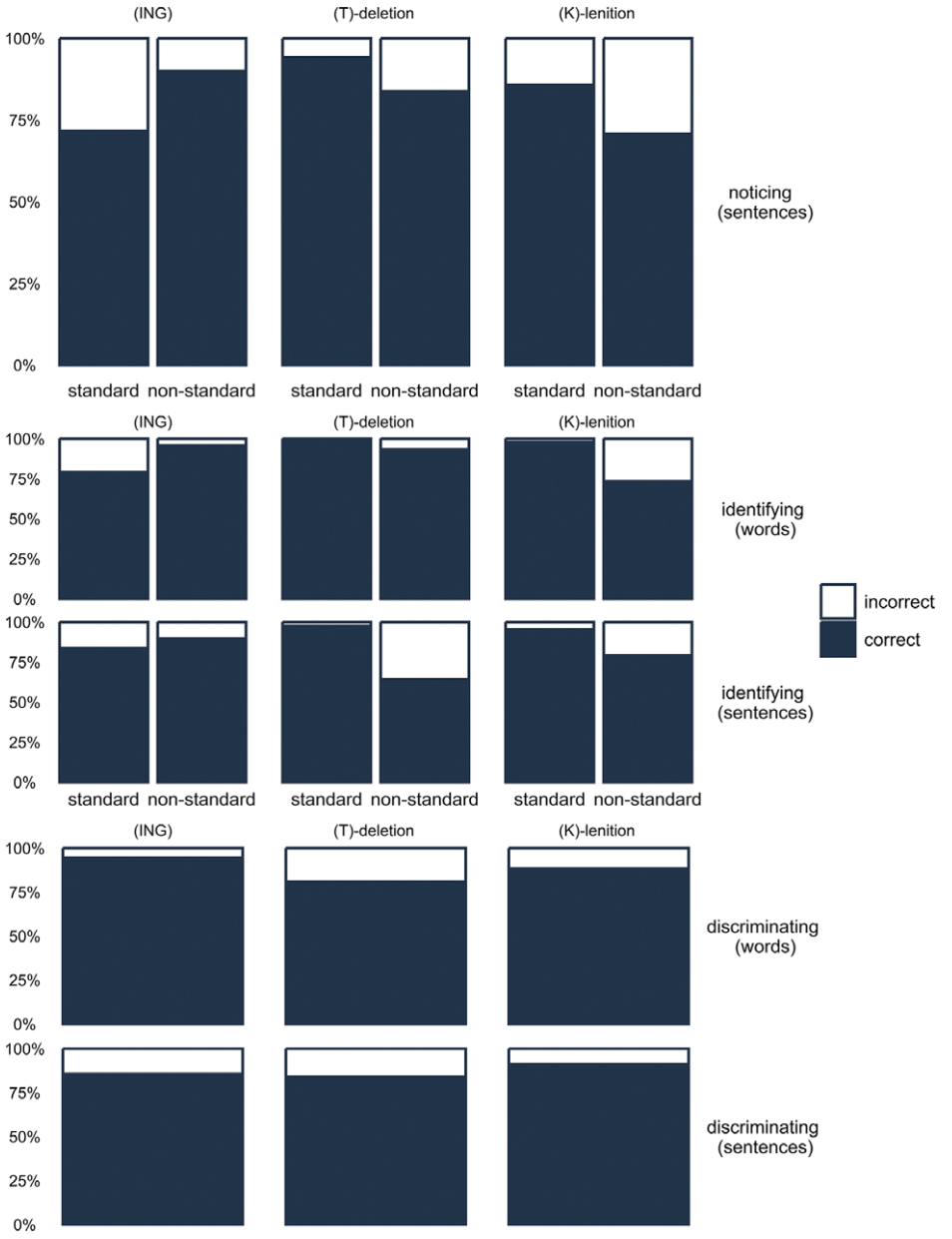

Figure 1. Visualisation of the proportion of correct responses per TASK TYPE, VARIABLE and VARIANT

class participants had significantly more difficulty *identifying* (ING)-variants correctly. There was also a significant interaction term between CARRIER and VARIANT indicating lower CORRECTNESS rates when participants heard word-based stimuli that contained the

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 7. *(ING). Results of the* noticing tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	2.52	0.29	8.61	<0.001	***
variant = standard	1.42	0.29	4.96	<0.001	***

Table 8. *(T)-deletion. Results of the* noticing tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	1.96	0.25	7.75	<0.001	***
variant = standard	1.15	0.33	3.51	<0.001	***
region = South	0.58	0.30	1.90	0.057	

Table 9. *(K)-lenition. Results of the* noticing tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	0.97	0.17	5.65	<0.001	***
variant = standard	0.97	0.25	3.93	<0.001	***

Table 10. *(ING). Results of the* identification tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	3.26	0.39	8.41	<0.001	***
carrier = words	1.10	0.48	2.30	0.021	*
variant = standard	0.60	0.30	2.02	0.043	*
class = WC	0.89	0.32	2.76	0.006	**
carrier = words : variant = standard	1.53	0.49	3.10	0.002	**

standard variant but sentence-based stimuli that contained the non-standard variant (est.=-1.53, p=0.002, table 10).

As for (T)-deletion, we find the same pattern for VARIANT that we observed with the *noticing tasks.* Participants were significantly better at *identifying* the standard variant [t] as opposed to non-standard deleted-[t] (est.=3.42, p<0.001, table 11). A significant main e ect for CARRIER suggests that, as with (ING)-based *identi cation tasks*, respondents were more likely to identify the (T)-deletion variants correctly if they were carried by a word rather than a sentence (est.=2.34, p<0.001, table 11). Results for

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 11. *(T)-deletion. Results of the* identification tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	0.67	0.18	3.68	<0.001	***
carrier = words	2.34	0.36	6.47	<0.001	***
variant = standard	3.42	0.45	7.60	<0.001	***

Table 12. *(K)-lenition. Results of the* identification tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	1.80	0.27	6.72	<0.001	***
carrier = words	0.47	0.33	1.41	0.159	
variant = standard	2.00	0.38	5.23	<0.001	***
carrier = words: variant = standard	1.22	0.62	1.96	0.050	*

the *identi cation tasks* of (K)-lenition variants indicate that [k] was more correctly identified than the non-standard [x] (est.=2.00, p<0.001, table 12). There was an interaction e ect between CARRIER and VARIANT pointing to significantly higher CORRECTNESS rates when participants heard word-based stimuli that contained the standard variant but sentence-based stimuli that contained the non-standard (est.=1.22, p=0.050, table 12).

Discrimination tasks. Respondents were generally better at *discriminating* between the variants of (ING) when they were presented with individual words as opposed to sentences (est.=1.15, p<0.001, table 13). This pattern contrasts with the findings for both (T)-deletion (est.=-0.22, p<0.204, table 14) and (K)-lenition (est.=-0.30, p=0.165, table 15), where CARRIER was not found a significant predictor of respondents CORRECTNESS rates. A correct *discrimination* of (ING)-variants was furthermore predicted by CLASS and REGION. WC participants (est.=-0.51, p=0.025, table 13) and Southerners (est.=-0.75, p=0.001, table 13) performed significantly worse in the *discrimination tasks* than MC respondents and participants from the North.

Table 13. *(ING). Results of the* discrimination tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	2.44	0.24	10.32	<0.001	***
carrier = words	1.15	0.24	4.71	<0.001	***
class = WC	0.51	0.23	2.23	0.025	*
region = south	0.75	0.23	3.23	0.001	**

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

Table 14. *(T)-deletion. Results of the* discrimination tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	1.66	0.12	13.31	<0.001	***
carrier = words	0.22	0.17	1.27	0.204	

Table 15. *(K)-lenition. Results of the* discrimination tasks *binomial logistic regression model*

	Estimate	Std error	z-value	p-value	
(Intercept)	2.34	0.16	14.51	<0.001	***
carrier = words	0.30	0.22	1.39	0.165	

5 Discussion and conclusion

In this study, we set out to answer three questions: first, we asked about the proportion of correct *noticing*, *identifying* and *discriminating* of variants of (ING), (T)-deletion and (K)-lenition at word- and sentence-level. Second, we sought to determine task-dependent factors and listener characteristics that *may* lead to incorrect *noticing*, *identifying* and *discriminating.* Finally, we asked if it would be worthwhile to include tasks that assess respondents ability to *notice*, *identify* and *discriminate* variable linguistic features in SES and, if so, which task type is most suitable.

RQ1. Exploring the extent of listeners ability to *notice*, *identify* and *discriminate* variable linguistic features, we found that, indeed, L1 English speakers have these abilities but do not *identify* and *discriminate* all variants correctly. Also, they show relatively low CORRECTNESS rates for *noticing.* In SES, we cannot assume that all listeners will correctly *notice*, *identify* and *discriminate* the VARIANTS of interest, which *may*, in turn, a ect listener evaluation. CORRECTNESS rates are also influenced by various factors that are relevant to SES.

RQ2. Regarding task-dependent and listener characteristics that function as predictors of respondents CORRECTNESS rates, we found VARIANT and CARRIER to show a systematic e ect. The e ect of VARIANT becomes particularly obvious with *noticing* and *identifying.* CORRECTNESS rates di er drastically among VARIANTS (table 6), and we observe two patterns: one for (ING) and a di erent one for (T)-deletion and (K)-lenition. For (ING), *non-standard* [n] is substantially more frequently *noticed* and *identi ed* correctly. Yet, for (T)-deletion and (K)-lenition, it is the *standard* variants that are correctly *noticed* and *identi ed* more frequently. Regression models show that this

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

di erence in CORRECTNESS rates is statistically significant for all variables tested in the *noticing* and *identi cation tasks*.[7]

One question remains: why do (ING) on the one hand and (T)-deletion and (K)-lenition on the other show contrasting patterns? Exploring these asymmetries was not part of our research design. Thus, we can only o er post-hoc explanations: regarding (i) variant frequency within a variety and (ii) the acoustic signal. Concerning (i), deleted-[t] and [x] generally occur less frequently than the standard variants do in the British varieties investigated; for (ING), this is not always the case, especially not in Northern and South-Western areas within England, where [n] is more frequent than [ŋ] (see section 2). More experience with a variant may translate into more successful *noticing*, *identifying* and *discriminating*. While we did find a main e ect for REGION in one of the models (i.e. Northern listeners were somewhat better in the *discrimination tasks* for (ING)), there were no interaction e ects for VARIANT and REGION in any of the models, which is what we would expect if variant frequency were indeed influencing CORRECTNESS rates. Thus, we are forced to dismiss this as an explanation.

This leaves explanation (ii). It may be inherently easier to *notice, identify* or *discriminate* some sounds rather than others in certain phonological contexts. For example, it may be easier to *notice, identify* or *discriminate* a sound that is present than the lack of a sound, such as deleted-[t], especially when the preceding sound is also alveolar (e.g. /n/). Finding evidence for this hypothesis is clearly a task for future research. Nonetheless, the asymmetric *noticing* and *identi cation* of variants is an important finding of our study that seems relevant to how SES are designed.

As for CARRIER as a predictor of respondents CORRECTNESS rates, we expected more correct answers when the feature of interest occurred in single words than when it occurred in a carrier sentence due to the additional cognitive load that processing a sentence is associated with. While some models support this assumption, findings are not as clear-cut as expected. For correct *identi cation* or *discrimination* of (K)-lenition variants, CARRIER does not matter. It also does not matter for the *discrimination* of (T)-deletion variants and for *identifying* [ŋ], where CORRECTNESS rates are higher for sentences rather than words. Thus, variants in words are not more correctly *noticed*, *identi ed* and *discriminated* in all cases. Besides cognitive load, other factors must be influencing CORRECTNESS rates in di erent CARRIER types, which should be explored in the future. We believe that the phonological context of variants, for instance, may be a confounding factor, which could explain the asymmetries in CARRIER type. In words, variants are followed by a pause; in sentences, they are followed by a vowel. For some variants, the following pause in a carrier word may reduce any advantages that a lower cognitive load in *identi cation* and *discrimination* may bring.[8]

[7] Thus, we can exclude the possibility that orthographic representation has biased the *noticing* and *identi cation* of variants substantially and can conclude that listeners really did pay attention to what they heard.

[8] As pointed out by one of the reviewers, another potential factor of influence may be actual di erences in the tokens.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

The e ect of TASK TY E is a further aspect to consider. As *discrimination tasks* (*same* vs *different*) do not allow for determining CORRECTNESS rates at VARIANT-level, a statistical comparison of methods is not possible. It also does not seem necessary considering our primary goal of uncovering the extent to which English L1 speakers correctly *notice, identify* or *discriminate* di erent variables. It is clear that, moving from *discrimination*, to *identi cation* and *noticing*, the tasks become increasingly more demanding (see table 1). This is reflected in our data: CORRECTNESS rates are comparatively high for the *discrimination tasks*, and results also reflect the additional cognitive load of the *noticing tasks* when compared to *identi cation* in the majority of cases. We created additional regression models to compare all six variants at sentence-level in the *noticing* and the *identi cation tasks.* Results show that *identi cation* rates are significantly more often correct than *noticing* rates for [ŋ], [t], [k] and [x] (p<.05 in all cases). There was no significant di erence for [n], and deleted-[t] is more often correctly *noticed* than *identi ed* (p<.001, see figure 2). These task-type asymmetries for di erent variants suggest that, in some cases (e.g. deleted-[t]), CORRECTNESS rates are influenced by more than just the cognitive demands of a particular task type. What factors these are remains a question for future research.

Returning to the task types, it is obvious that, in the majority of cases, results for task type di er because they seem to test di erent processes. *Discrimination tasks* assess a participant s ability to distinguish between pronunciation variants *indirectly* (*same* vs *different*). *Identi cation tasks*, in turn, also test if participants can hear variants correctly but pose the greater challenge of hearing for certain variants. In addition to hearing the variants of a variable correctly, *noticing tasks* rely on variants standing out to listeners. A direct comparison of results for *noticing* and *identifying* (see figure 2) makes apparent that these two task types test di erent processes as CORRECTNESS rates for almost all variants di er clearly between both task types.

RQ3. Our results also provide us with some guidance as to whether *noticing, identi cation* and *discrimination tasks* should be included in SES in any way. As CORRECTNESS rates fall into the mid-60s for some variants, testing whether they

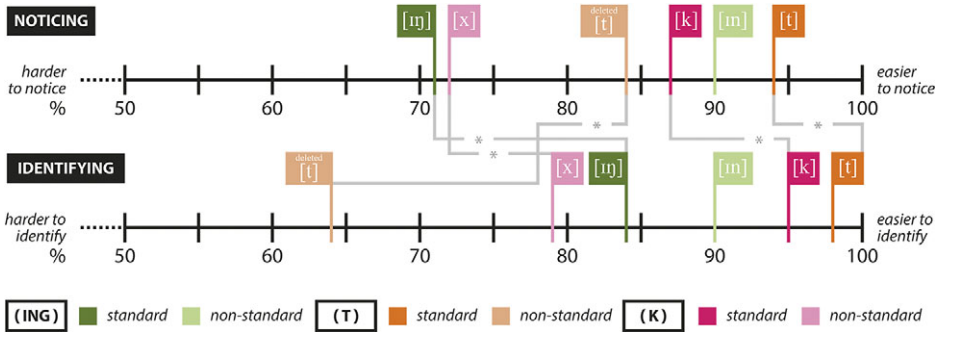

Figure 2. Comparative overview of results of the *noticing* and *identi cation tasks*
* = Di erence is statistically significant.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

improve models in SES may certainly be worthwhile. What measure would one collect then in addition to data from actual SES? Considering that CORRECTNESS rates di er for variants, *discrimination tasks* are not a suitable task type to include as they do not provide any variant-specific response data due to the *same* versus *different* answer options. *Noticing* and *identi cation tasks*, on the other hand, provide these data. Of these two, *identi cation tasks* seem to be most suitable to include in an SES since this task type is somewhat easier than *noticing tasks.* When implemented towards the end of a survey, which is the most likely scenario, they might be more e ective when listener fatigue tends to become a serious issue. Also, in *identi cation tasks*, carrier sentences seem desirable as they are more in line with guises typically used in SES and they are not necessarily more challenging test materials than words. Using carrier sentences can also provide more control and variety regarding the element that follows the target feature than carrier words can. Generally speaking, *identi cation tasks* seem particularly suitable for studies that seek to assess the actual ability of a listener to hear sounds and label them correctly. To what extent the evaluation of language is influenced by the degree to which di erent listeners can *identify* features remains an open question. *Noticing tasks*, too, could in principle be used as follow-up tasks to an SES. They assess both the ability to hear a variant and the degree to which a given variant stands out. Regardless of their *noticing* abilities, participants may still evaluate guises however, they may do so di erently. Also, *noticing tasks* should be implemented if naturalness is of high priority.

Finally, our results provide some guidance on what variants to include in *identi cation* or *noticing tasks*, namely those variants that seem to trouble participants and not those with CORRECTNESS rates of more than 90 per cent, as these would not be diagnostic of any *identi cation* or *noticing* issues. Future research needs to establish whether the inclusion of CORRECTNESS rates for individual participants does improve speaker evaluation models at all, or whether other factors provide better measures. We tested some social characteristics of listeners and did not find much evidence that they influence CORRECTNESS rates. Cognitive factors, however, might do so, such as the capacity of participants short-term as well as simple and complex working memory, or psychometric measures, for instance, elicited through the BAPQ (Hurley *et al.* 2007) or the Self-Monitoring Questionnaire (Cramer & Gruman 2002).

Our results also speak to sociolinguistic theory. If social meanings are indexed by variants (see Campbell-Kibler 2011) and if the ability to *notice* these varies (see our results), then it follows that the social meanings of some variants *may* be missed by *some* speakers. Alternatively, those listeners with high CORRECTNESS rates in *noticing* variants may evaluate di erently than those with low CORRECTNESS rates. In any case, a poor performance at *noticing tasks* does not mean that listeners will not evaluate these features.

Of course, this reasoning is based on the assumption that individual variants *are* noticed in the perception of social meanings. Alternatively, listeners may access social meanings based on perceived styles and co-occurring features (Levon 2014; Pharao & Maegaard 2017; Schleef 2023). Still, even when listeners focus on styles, the issue

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

remains that some variants are more readily *noticed*, *identi ed* and *discriminated* than others. And indeed, research that explores co-occurring variants has found that some features are trumped by others, which has normally been explained with a view to social salience (e.g. Levon 2014; Schleef 2023). Since we pointed listeners directly to the variable context, we assume that social salience did not influence CORRECTNESS rates to any large extent (especially so in the *identi cation* and *discrimination tasks*). Instead, we have argued that CORRECTNESS rates may be impacted by the perceptibility of a sound in a specific linguistic context.

In conclusion, this study provides more evidence for the disjunct of production and perception in variationist sociolinguistics (Levon 2014; Schleef 2023). While research on style production has argued that features cluster into styles to create social meanings (Eckert 2012), it is by no means clear that listeners can perceive all features equally and recognise the intended meanings. The mechanisms involved in *noticing*, *identifying* and *discriminating* and in processing the social meanings of individual or co-occurring features remain rather obscure. The present study raises questions regarding the forms that can be *noticed*, *identi ed* and *discriminated*, and by whom, and results suggest that a more rigorous elicitation of such information might be useful for a more complete understanding of the nature of social evaluation.

Authors address:

Department of English and American Studies
University of Salzburg
Erzabt-Klotz-Straße 1
5020 Salzburg
Austria
jana.pflaeging@plus.ac.at (corresponding author)

References

Bailey, George. 2019. Emerging from below the social radar: Incipient evaluation in the North West of England. *Journal of Sociolinguistics* 23(1), 3 28.

Baranowski, Maciej & Danielle Turton. 2020. TD-deletion in British English: New evidence for the long-lost morphological e ect. *Language Variation and Change* 32(1), 1 23. https://doi.org/10.1017/S0954394520000034

Bassili, John N. & Rick D. Brown. 2005. Implicit and explicit attitudes: Research, challenges, and theory. In Dolores Albarracin, Blair T. Johnson & Mark P. Zanna (eds.), *The handbook of attitudes*, 543 74. Mahwah, NJ: Lawrence Erlbaum Associates.

Bayley, Robert. 1994. Consonant cluster reduction in Tejano English. *Language Variation and Change* 6(3), 303 26. https://doi.org/10.1017/S0954394500001708

Bender, Emily M. 2005. On the boundaries of linguistic competence: Matched-guise experiments as evidence of knowledge of grammar. *Lingua* 115(11), 1579 98. https://doi.org/10.1016/j.lingua.2004.07.005

Boswijk, Vincent & Matt Coler. 2020. What is salience? *Open Linguistics* 6, 713 22.

Brooks, Mollie E., Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler & Benjamin M. Bolker. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal* 9(2), 378 400. https://doi.org/10.32614/RJ-2017-066
Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech* 82, 32 64. https://doi.org/10.1215/00031283-2007-002
Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21, 135 56. https://doi.org/10.1017/S0954394509000052
Campbell-Kibler, Kathryn. 2011. The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change* 22, 423 41. https://doi.org/10.1017/S0954394510000177
Coetzee, Andries W. & Shigeto Kawahara. 2013. Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 31(1), 47 89.
Cooper, Robert L. 1975. Introduction to language attitudes II. *International Journal of the Sociology of Language* 6, 5 9.
Cramer, Kenneth M. & Jamie A. Gruman. 2002. The Lennox and Wolfe Revised Self-Monitoring Scale: Latent structure and gender invariance. *Personality and Individual Differences* 32, 627 37.
De, Zachary. 2017. Race, gender, and /u/: Social perceptions of a non-stereotype feature. MA thesis, The Ohio State University.
Drager, Katie. 2018. *Experimental research methods in Sociolinguistics*. London: Bloomsbury.
Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41, 87 100. https://doi.org/10.1146/annurev-anthro-092611-145828
Fasold, Ralph W. 1972. *Tense marking in Black English: A linguistic and social analysis*. Arlington, VA: Center for Applied Linguistics.
Grover, Vikas, Valerie L. Shafer, Luca Campanelli, D. H. Whalen & Erika S. Levy. 2021. Perception of American English consonants /v/ and /w/ by Hindi speakers of English. *Journal of Second Language Pronunciation* 7(3), 370 407. https://doi.org/10.1075/jslp.20033.gro
Guy, Gregory R. 1980. Variation in the group and the individual: The case of final stop deletion. In William Labov (ed.), *Locating language in time and space*, 1 36. New York: Academic Press.
Guy, Gregory R. 1991. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3(1), 1 22. https://doi.org/10.1017/S0954394500000429
Hay, Jennifer, Paul Warren & Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34, 458 84.
Hazen, Kirk. 2006. IN/ING variable. In Keith Brown (ed.), *Encyclopaedia of language and linguistics*, vol. 5, 581 4. Amsterdam: Elsevier. https://doi.org/10.1016/B0-08-044854-2/04716-7
Hazen, Kirk. 2011. Flying high above the social radar: Coronal stop deletion in modern Appalachia. *Language Variation and Change* 23(1), 105 37. https://doi.org/10.1017/S0954394510000220
Hurley, Robert S. E., Molly Losh, Morgan Parlier, J. Steven Reznick & Joseph Piven. 2007. The Broad Autism Phenotype Questionnaire. *Journal of Autism and Developmental Disorders* 37, 1679 90.
Kerswill, Paul & Ann Williams. 2002. 'Salience as an explanatory factor in language change: Evidence from dialect levelling in urban England. In Mari C. Jones & Edith Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors*, 81 110. Berlin and Boston: De Gruyter.
Houston, Ann C. 1985. Continuity and change in English morphology: The variable (ING). PhD dissertation, University of Pennsylvania.

Kristiansen, Gitte. 2008. Style-shifting and shifting styles: A socio-cognitive approach to lectal variation. In Gitte Kristiansen & René Dirven (eds.), *Cognitive Sociolinguistics: Language variation, cultural models, social systems*, 45 90. Berlin and Boston: De Gruyter.
Labov, William. 1989. The child as a linguistic historian. *Language Variation and Change* 1, 85 97.
Labov, William. 2001. *Principles of linguistic change*, vol. 2: *Social factors.* Oxford: Blackwell.
Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski & Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15(4), 431 63. https://doi.org/10.1111/j.1467-9841.2011.00504.x
Lambert, W. E., R. C. Hodgson, R. C. Gardner & S. Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60(1), 44 51. https://doi.org/10.1037/h0044430
Levon, Erez. 2014. Categories, stereotypes and the linguistic perception of sexuality. *Language in Society* 43, 539 66.
Levon, Erez & Isabelle Buchstaller. 2015. Perception, cognition, and linguistic structure: The e ect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27, 319 48. https://doi.org/10.1017/S0954394515000149
Levon, Erez & Sue Fox. 2014. Social salience and the sociolinguistic monitor: A case of ING and TH-fronting in Britain. *Journal of English Linguistics* 42(3), 185 217.
LimeSurvey GmbH. 2023. *LimeSurvey: An open source survey tool* [computer software]. Hamburg, Germany. Available at www.limesurvey.org
Mackay, Bradley. 2024. The heterogeneity of sex-based social categories: A quantitative analysis of voice onset time and T-glottaling in gay and straight speech. PhD thesis, University of Salzburg, Austria.
Montgomery, Chris & Emma Moore. 2018. Evaluating S(c)illy voices: The e ects of salience, stereotypes, and co-present language variables on real-time reactions to regional speech. *Language* 94(3), 629 61.
Nycz, Jennifer. 2016. Awareness and acquisition of new dialect features. In Anna Babel (ed.), *Awareness and control in sociolinguistic research*, 62 79. Cambridge: Cambridge University Press.
Osgood, Charles E., George J. Suci & Percy H. Tannenbaum. 1957. *The measurement of meaning*. Chicago: University of Illinois Press.
Pavlík, Radoslav. 2017. Some new ways of modelling t/d deletion in English. *Journal of English Linguistics* 45(3), 195 228. https://doi.org/10.1177/0075424217712358
Pharao, Nicolai & Marie Maegaard. 2017. On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics* 55, 1141 67. https://doi.org/10.1515/ling-2017-0023
Pharao, Nicolai, Marie Maegaard, Janus Spindler Møller & Tore Kristiansen. 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in di erent prosodic contexts. *Language in Society* 43, 1 31. https://doi.org/10.1017/S0047404513000857
Preston, Dennis. 2010. Variation in language regard. In Peter Gilles, Joachim Scharloth & Evelyn Zeigler (eds.), *Variatio Delectat: Empirische Evidenzen und theoretische Passungen sprachlicher Variation*, 7 27. Frankfurt/Main: Peter Lang.
Preston, Dennis. 2017. The cognitive foundations of language regard. *Poznan Studies in Contemporary Linguistics* 53(1), 17 42.
Prolific. 2023. London. Available at www.prolific.com
Purschke, Christoph. 2011. *Regionalsprache und Hörerurteil*. Stuttgart: Franz Steiner Verlag.

R Core Team. 2021. *R: A language and environment for statistical computing* [computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available at www.R-project.org
Santa Ana, Otto. 1996. Sonority and syllable structure in Chicano English. *Language Variation and Change* 8(1), 63 89. https://doi.org/10.1017/S0954394500001071
Schleef, Erik. 2023. Mechanisms of meaning making in the co-occurrence of pragmatic markers with silent pauses. *Language in Society* 52(1), 79 105.
Schleef, Erik & Nicholas Flynn. 2015. Ageing meanings of (ing): Age and indexicality in Manchester, England. *English World-Wide* 36, 47 89.
Schleef, Erik, Nicholas Flynn & Will Barras. 2017. Regional diversity in social perceptions of (ing). *Language Variation and Change* 29, 29 56.
Schleef, Erik, Miriam Meyerho & Lynn Clark. 2011. Teenagers acquisition of variation: A comparison of locally-born and migrant teens realisation of English (ing) in Edinburgh and London. *English World-Wide* 32(2), 206 36.
Silverstein, Michael. 2003. Indexical order and the dialects of sociolinguistic life. *Language Communication* 23, 193 229.
Smith, Jennifer, Mercedes Durham & Liane Fortune. 2009. Universal and dialect-specific pathways of acquisition: Caregivers, children, and t/d deletion. *Language Variation and Change* 21(1), 69 95. https://doi.org/10.1017/S0954394509000039
Squires, Lauren. 2014. Talker specificity and the perception of grammatical variation. *Language, Cognition and Neuroscience* 29(7), 856 76.
Tagliamonte, Sali A. 2004. Somethi[ŋ] s goi[n] on Variable (ing) at Ground Zero. In Britt-Louise Gunnarsson, Lena Bergstrom, Gerd Eklund, Sta an Fidell, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundergren & Mats Thelander (eds.), *Language variation in Europe: Papers from the Second International Conference on Language Variation in Europe ICLaVE2, Uppsala, Sweden, 12th 14th June 2003*, 390 403. Uppsala: Uppsala University.
Tagliamonte, Sali A. & Rosalind Temple. 2005. New perspectives on an ol variable: (t, d) in British English. *Language Variation and Change* 17(3), 281 302. https://doi.org/10.1017/S0954394505050118
Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
Trudgill, Peter. 1986. *Dialects in contact*. London: Blackwell.
Vaughn, Charlotte. 2022a. What carries greater social weight, a linguistic variant s local use or its typical use? *Proceedings of the Annual Meeting of the Cognitive Science Society* 44(44), 513 20.
Vaughn, Charlotte. 2022b. The role of internal constraints and stylistic congruence on a variant s social impact. *Language Variation and Change* 34(3), 331 54. https://doi.org/10.1017/S0954394522000175
Wagner, Susan E. & Ashley Hesson. 2014. Individual sensitivity to the frequency of socially meaningful linguistic cues a ects language attitudes. *Journal of Language and Social Psychology* 33(6), 651 66. https://doi.org/10.1177/0261927X14528713
Watson, Kevin. 2006. Phonological resistance and innovation in the North-West of England. *English Today* 22(2), 55 61.
Watson, Kevin. 2007. Liverpool English. *Journal of the International Phonetics Association* 37(3), 351 60. https://doi.org/10.1017/S0025100307003180
Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
Wolfram, Walt A. 1969. *A sociolinguistic description of Detroit Negro speech*. Washington, DC: Center for Applied Linguistics.
Zarcone, Alessandra, Marten van Schijndel, Jorrig Vogels & Vera Demberg. 2016. Salience and attention in surprisal-based accounts of language processing. *Frontiers in Psychology* 7, 844. https://doi.org/10.3389/fpsyg.2016.00844

Appendix

A.1 Test stimuli

Noticing tasks

The following stimuli (sorted by variable) were used in the *noticing tasks* (sentences only).

- He was **running** [ŋ][n] out of excuses not to go. (1st block)
- They were **making** [ŋ][n] excellent progress on the project. (1st block)
- She was **reading** [ŋ][n] about current a airs over breakfast. (1st block)
- She was **writing** [ŋ][n] a note when her pen ran out of ink. (2nd block)
- He was **talking** [ŋ][n] angrily on the phone when he came in. (2nd block)
- She was **walking** [ŋ][n] along the road when the tree fell. (2nd block)

- The undercover **agent** [t][ ] infiltrated the gang. (1st block)
- His boss s positive **comment** [t][ ] on his work made his day. (1st block)
- The boy s **present** [t][ ] is waiting for him under the tree. (1st block)
- The happy **client** [t][ ] immediately placed an order. (2nd block)
- In a **moment** [t][ ] of weakness he admitted his crimes. (2nd block)
- He was a **student** [t][ ] at a prestigious boarding school. (2nd block)

- She hung a **hammock** [k][x] in the shade beneath the trees. (1st block)
- He carries **sunblock** [k][x] in his bag when he goes hiking. (1st block)
- The big **padlock** [k][x] on the garage door was cut open. (1st block)
- In her **notebook** [k][x] Anne has a list of films she wants to see. (2nd block)
- She enjoyed the **outlook** [k][x] over the surrounding fields. (2nd block)
- They had a **food truck** [k][x] at their wedding reception. (2nd block)

- The **cat/car** is behind the house. [filler]
- Can you see the little **mouse/house**? [filler]
- He heard the **rain/train** through the window. [filler]

Identi cation tasks

The following stimuli (sorted by variable) were used in the sentence-based *identi cation tasks.* The focus words (in **bold**, e.g., *running, making, reading, agent, etc.*) were also used as part of the word-based *noticing tasks.* To save space, the list of single words is not repeated here.

- She was **running** [ŋ][n] in order to catch the bus.
- He was **making** [ŋ][n] a sandwich when his mother walked in.
- He was **reading** [ŋ][n] a novel while she watched TV.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

- He was **writing** [ŋ][n] a letter to his favourite aunt.
- They were **talking** [ŋ][n] about local politics.
- I was **walking** [ŋ][n] in the park when I saw the fox.

- The young **agent** [t][ ] enthusiastically pitched the role.
- The insulting **comment** [t][ ] online worsened his mood.
- The wrapped **present** [t][ ] is on the table by the cake.
- The new **client** [t][ ] expressed satisfaction with the idea.
- She took a **moment** [t][ ] and relaxed on the sofa.
- The hard-working **student** [t][ ] excelled in her studies.

- The colourful **hammock** [k][x] is swaying in the breeze.
- She applied **sunblock** [k][x] once she got to the beach.
- She uses a **padlock** [k][x] and chain to secure her bike.
- He got his **notebook** [k][x] and took down the statement.
- He gave an **outlook** [k][x] on future plans and projects.
- The new **food truck** [k][x] is becoming very popular.

Discrimination tasks

The following stimuli (sorted by variable) were used in the sentence-based *discrimination tasks*, for which a combination of two sentences (same wording) with focus words were used. The focus words (in **bold**, e.g., *running, making, reading, agent, etc.*) were also used as part of the word-based *discrimination tasks.* To save space, the list of single words is not repeated here.

- She was **running** [ŋ][n] in order to catch the bus.
- He was **making** [ŋ][n] a sandwich when his mother walked in.
- He was **reading** [ŋ][n] a novel while she watched TV.
- He was **writing** [ŋ][n] a letter to his favourite aunt.
- They were **talking** [ŋ][n] about local politics.
- I was **walking** [ŋ][n] in the park when I saw the fox.

- The young **agent** [t][ ] enthusiastically pitched the role.
- The insulting **comment** [t][ ] online worsened his mood.
- The wrapped **present** [t][ ] is on the table by the cake.
- The new **client** [t][ ] expressed satisfaction with the idea.
- She took a **moment** [t][ ] and relaxed on the sofa.
- The hard-working **student** [t][ ] excelled in her studies.

- The colourful **hammock** [k][x] is swaying in the breeze.
- She applied **sunblock** [k][x] once she got to the beach.
- She uses a **padlock** [k][x] and chain to secure her bike.

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

- He got his **notebook** [k][x] and took down the statement.
- He gave an **outlook** [k][x] on future plans and projects.
- The new **food truck** [k][x] is becoming very popular.

A.2 Overview of the survey

- **Proli c ID:** Please enter your Prolific ID into the text box.
- **Training Page:** In this survey, you will listen to and answer questions about audio recordings. Please take a moment now to make sure you are in a quiet environment. Use the following audio samples to ensure your browser can play audio files and to adjust the volume to a comfortable level. If you are unable to hear the audio clearly due to technical difficulties or any other reason, such as a hearing impairment, please do not continue with the survey.

  In this survey, you will be asked to decide whether certain words were pronounced in specific ways. For example, you might be asked to decide whether you heard a *k* sound in the word *book* or another sound that sounded more like [x] (like the final sound in *loch* in Scottish English). We use the symbol [x] for the second option. To help you understand what we mean, please listen to the audio examples below as often as you like.

  k as in *book* Recording
  [x] as in *boo*[x] Recording

  You might also be asked to decide whether or not you heard a *t* sound at the end of a word. To help you understand what is meant, please listen to the audio examples below as often as you like.

  -nt as in *brilliant* Recording
  -n as in *brillian* Recording

  Or you might be asked to decide whether you heard an *-ing* or an *-in* sound at the end of a word. To help you understand what is meant, please listen to the audio examples below as often as you like.

  -ing as in *ghting* Recording
  -in as in *ghtin* Recording

- **PART 1. Introduction:** This is Part 1 of the survey. In this part, you will be asked to, first, listen to an audio recording. Then, you will be given a question with several answer options. Click *Play* to start the recording. Once the recording has played through, the audio element will disappear. Then, answer the question that appears and decide on one of the answer options. Remember to choose the answer that you

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

actually hear, rather than what you expect to hear based on Standard English. Now, click *Next* to practise with a trial question.

- **PART 1. Trial question:** Click *Play* to start the recording. Once the recording has played through, the audio element will disappear. Now, answer the question that appears and decide on one of the answer options.

  Recording

  You have just heard a recording of the word *went.* How was it pronounced?

  A: went
  B: vent
  C: I don t know.

- **PART 1. Noticing tasks** (in randomised order)
  - 3 **x (ING)**
  - **3 x (T)-deletion**
  - **3 x (K)-lenition**
  - **1 x Attention Check**
  - **3 x Fillers**

- **PART 2. Introduction:** Well done You ve finished the first part. This is Part 2 of the survey. In this part, a question with several answer options will be shown to you. You will also be asked to listen to an audio recording. Click *Play* to start the recording. Once the recording has played through, the audio element will disappear. Remember to choose the answer that you actually hear, rather than what you expect to hear based on Standard English. Now, click *Next* to practise with a trial question.

- **PART 2. Trial question:** Click *Play* to start the recording. Once the recording has played through, the audio element will disappear. Then, answer the question given. Remember to choose the answer that you actually hear, rather than what you expect to hear based on Standard English.

  You will hear a recording of the word *went.*

  Recording

  How was it pronounced?
  - went
  - vent
  - I don t know.

- **PART 2. Identi cation tasks** (in randomised order)
  - **6 x (ING)**
  - **6 x (T)-deletion**
  - **6 x (K)-lenition**
  - **1 x Attention Check**

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

- **PART 3. Introduction:** Well done You ve finished the second part. This is Part 3 of the survey. In this part, you will be asked to listen to a recording of two words. Also, a question with several answer options will be shown to you. Click *Play* to start the recording. Once the recording has played through, the audio element will disappear. Then, answer the question. Remember to choose the answer that you actually hear, rather than what you expect to hear based on Standard English. Now, click *Next* to practise with a trial question.

- **PART 3. Trial question:** Click *Play* to start the recording. Once the recording has played through, the audio element will disappear. Remember to choose the answer that you actually hear, rather than what you expect to hear based on Standard English.

  Focus on the initial sound in the word *volunteers* in both recordings. Click SAME if you think the initial sound was the same both times. Click DIFFERENT if you think it sounded di erent.

  Recording

  How was it pronounced?
  - SAME
  - DIFFERENT
  - I don t know.

- **PART 2. Discrimination tasks** (in randomised order)
  - **6 x (ING)**
  - **6 x (T)-deletion**
  - **6 x (K)-lenition**
  - **1 x Attention Check**

- **PART 4. Biographical questions**
- **Gender:** Which of the following best describes your gender identity?
  - woman
  - man
  - non-binary
  - prefer not to say
  - Other: [*textbox*]
- **Age:** How old are you? [*textbox, numbers only*]
- **Class:** What social class do you think you belong to?
  - lower working class
  - upper working class
  - lower middle class
  - upper middle class
  - lower upper class
  - upper upper class
- **L1:** Is English your native/first language?
  - Yes
  - No

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press

- **L1 If 'No :**
  What language would you consider your native/first language? [*textbox*]
  How many years have you lived in the UK? [*textbox, numbers only*]

- **Place of residence:** What is your current place of residence?
  - North East, England (Tees Valley, Durham, Northumberland and Tyne and Wear)
  - North West, England (Cumbria, Greater Manchester, Lancashire, Merseyside)
  - Yorkshire and the Humber, England (East Riding, North Lincolnshire and Yorkshire)
  - London, England
  - South East, England (Berkshire, Buckinghamshire, Oxfordshire, Surrey, Sussex, Kent, Hampshire and Isle of Wight)
  - South West, England (Gloucestershire, Wiltshire and Bristol/Bath area, Dorset and Somerset, Cornwall and Isles of Scilly, Devon)
- **Postcode:** Please provide the first three characters of your UK postcode. [*textbox*]
- **Audio:** What type of audio output device did you use for this study?
  - headphones
  - in-built computer speakers
  - external speakers

https://doi.org/10.1017/S1360674324000625 Published online by Cambridge University Press