Studying individual differences in the social environment to better understand language learning and processing

Abstract: People learn language from their social environment. Therefore, individual differences in the input that their social environment provides could influence their linguistic performance. Nevertheless, investigation of the role of individual differences in input on performance has been mostly restricted to first and second language acquisition. In this paper I argue that individual differences in input can influence linguistic performance even in adult native speakers. Specifically, differences in input can affect performance by influencing people’s knowledgebase, by modulating their processing manner, and by shaping expectations. Therefore, studying the role that individual differences in input play can improve our understanding of how language is learned, processed and represented.

Keywords: individual differences, social networks, PsychLingvar

1 From noise to structured differences

The realization that individual differences in performance can be informative about language processing and representation and are not merely noise that needs to be reduced has only gained popularity in Psycholinguistics over the last couple of decades. For many decades, experimental studies in psycholinguistics, as well as in other fields, focused on the commonalities, on how people in general perform certain tasks, and how the average participant is influenced by certain manipulations. Individual differences in performance were treated as noise. After all, by definition, participants’ score on any task is composed of their “true” score and the measurement error, and so differences might simply be noise. If the experiment is conducted well, the measurement errors should cancel each other, and the general trend would be discovered. It is only in the 90s that the systematicity in this “noise” has started to be explored and integrated into models of language processing.

By far the most studied individual differences in psycholinguistics are those related to executive function. Already in the 70s, models of language processing often assigned memory a role (e.g. Frazier
and Fodor 1978; Kintsch and van Dijk 1978), but it is mostly during the 90s that individual differences in memory and other aspects of executive function have started to be examined. The main line of research on this topic focused on individual differences in ability to hold an item in working memory, an ability that has been linked to success at text comprehension (Haarmann et al. 2003), ambiguity resolution (e.g. Miyake et al. 1994), low versus high attachment preference (Swets et al. 2007), and foreign language learning (e.g. Baddeley et al. 1998). Other lines of research focused on individual differences in inhibition, linking it to success at language acquisition (Darcy et al. 2014), degree of competitor activation across and within languages (Festman et al. 2010; Gollan et al. 2011; Linck et al. 2011; Lev-Ari and Keysar 2014; Lev-Ari and Peperkamp 2013, 2014), and success at inhibiting discourse-irrelevant information in either literal (e.g. Gernsbacher and Robertson 1995; Gernsbacher et al. 1990) or figurative language (e.g. Chiappe and Chiappe 2007; Kintsch 2001). Working memory has also been shown to influence the ability to integrate different types and sources of information online (Dagerman et al. 2006; Federmeier and Kutas 2005; Just and Carpenter 1992; Madden and Zwaan 2006; Pearlmutter and McDonald 1995; Traxler et al. 2005).

Different components of executive function, then, have been shown to be involved in learning and comprehension at different linguistic levels, and furthermore, individual differences along these executive functions have been shown to influence manner of processing as well as final success at comprehension and acquisition.

2 The role of individual differences in input

Unlike the case of cognitive abilities, individual differences in the social environment, and consequently, in the received input among adult native speakers have rarely been investigated. Initially, the role of input has mostly been discussed in the linguistic and psycholinguistic literature with regard to the consequences of lack of input. For example, cases such as Genie, who was raised deprived of linguistic input, have been presented as arguments for a critical age, above which one cannot acquire language with the same success (Curtiss et al. 1974). More recent studies compare deaf individuals whose age of first exposure to (sign) language varies, and show that the earlier the exposure, the better the linguistic outcome (Newport 1991).

In the developmental literature, amount of input and the Mean Length Utterances in caretakers’ speech have been shown to correlate with vocabulary growth and syntactic skills (Hoff 2003; Huttenlocher et al. 1991). Furthermore, using a natural manipulation, school periods versus summer vacation periods, it has been shown that children’s linguistic performance shows greater development over the school year than over an identical time period that includes a vacation, even though non-linguistic abilities, such as associative memory, show identical development across these time periods (Huttenlocher 1998). The amount and quality of the input children receive, then, seems to predict their linguistic development.

Input can differ, however, in more than just quantity, and, correspondingly, differences in input can influence linguistic performance in several ways. In this paper, I will focus on three ways in which input shapes behavior, such that individual differences in input can lead to differences in linguistic performance even among adult native speakers. First, I will describe perhaps the most direct and most investigated way in which input shapes performance, by influencing the knowledgebase that an individual has. If the input lacks information about certain uses or is unrepresentative of the language of the community at large, then interpretations based on its patterns might be inaccurate. Second, I will describe briefly how the properties of the input might influence what one attends to and the processing mechanisms one relies on during language processing. As with the former case, if the cues that are most relevant for analyzing the previously received input are different from those that are most informative in general, attention might be misdirected, and performance would be suboptimal. Lastly, I will describe how the input shapes our expectations, and how these influence our linguistic performance. Together, these sections will show how our success at
interpreting incoming input is shaped by the previous input we have received, and how, consequently, systematic differences in the input can lead to systematic differences in interpretation.

### 2.1 The role of input in shaping our knowledgebase

Our linguistic knowledge is based on the input we receive from the environment. An infant raised in Italy will acquire the vocabulary of Italian and not of Japanese, and unless exposed to physics or linguistic literature, an English speaker might never acquire the terms *white dwarf* or *fundamental frequency*, respectively. Similarly, if children are exposed to a dialect that doesn’t mark plural reliably, they will be less likely to produce and comprehend plural morphology (Miller and Schmitt 2010). Importantly, linguistic units are not independent of one another but form a system. Therefore, absence of specific forms can influence not only their acquisition, but also the acquisition of forms that contrast with them or are related to them. In other words, for linguistic input to be most useful, it needs to have the same structure as the language in the speech community at large.

A particularly nice demonstration of this can be seen in a study by Maye et al. (2002). In that study, 6 and 8-months-old infants were exposed to either a unimodal or a bimodal distribution of tokens on the /d/-/t/ continuum. Infants exposed to a bimodal distribution formed two phoneme categories. In contrast, infants exposed to a unimodal distribution formed a single phoneme category (Maye et al. 2002 but see Cristià et al. 2011). In other words, in order to know how to interpret input, we need to not only be exposed to all forms, but also to their appropriate relative distribution.

Similarly, the shape of the distribution influences whether our perceptual system is sensitive to changes along a dimension, leading us to be more sensitive to differences that matter in our language compared to those that do not. Thus, hearing a deviant sound in a stream of canonical sounds elicits a larger mismatch negativity (MNN), an ERP component sensitive to deviation, if the sound is classified as a different phoneme in our language than if it is not (Naatanen et al. 1997).

Another important aspect of the input is the number of sources from which it is generated, i.e., the number of speakers. Studies that manipulated this variable often found that acquisition is facilitated by exposure to multiple speakers compared with a single speaker. Thus 14-month old infants fail at understanding that the nonsense words, *puk* and *buk*, refer to different objects if they were earlier exposed to productions of one of these words by only one speaker. In contrast, if exposed to the same number of tokens of this word, but from 18 different speakers, infants succeed at mapping those words to different objects, indicating that they have acquired the phonemic contrast between /p/ and /b/ (Rost and McMurray 2009). Similar results have also been found with adults acquiring a new phonological variety. Thus, Japanese learners of English are better at learning the /l/-/ɹ/ contrast if trained with recordings from five speakers rather than only one (Lively et al. 1993), native speakers of American English adapt better to Chinese-accented English if exposed to four rather than one Chinese-accented speaker even when amount of input is matched (Bradlow and Bent 2008), and people are better at classifying new dialects if exposed to three rather than one speaker of each dialect (Clopper and Pisoni 2004).

Most of the studies that I have reviewed so far show that the structure of the input influences learning, either in infants acquiring their first language or in adults acquiring a new variety. Language learning, however, does not stop at infancy or childhood, but continues throughout our lives. We constantly sample the input in our environment and adjust our representations in accordance with it. Harrington et al. (2000) nicely demonstrated how linguistic representations change with time by analyzing Queen Elisabeth II’s pronunciation in her Christmas addresses over the years. They found that her pronunciation, much like the pronunciation of most British people, had shifted over the years, such that even the queen no longer speaks the Queen’s English in the strictest sense. Needless to say, similar changes can be found at other linguistic levels, as people might change their lexical choices and even their syntactic choices with time in response to changes in their linguistic input. Therefore, just like input influence language learning of new varieties, it might influence our skills as adult native speakers.
A couple of recent studies focusing on people’s social network provide some evidence that this is indeed the case. Social networks were defined as the people individuals regularly interact with, and so network members are not restricted to friends and family members, but include clients, colleagues, neighbors, and so forth. In one of these studies native Dutch speakers reported all their interactions for a period of 1 week. Then they were invited to the lab for a couple of phonological tasks, including a test of their ability to recognize vowels in Dutch pseudo-words that were embedded in noise. As the study used pseudo-words, participants could not rely on vocabulary knowledge but only on their phonological processing. Results indicated that participants who regularly interact with more people are better at understanding the speech in noise (Lev-Ari 2016). Furthermore, all participants were tested on a host of cognitive abilities, and results indicated both that cognitive abilities do not co-vary with social network size, and that the effect of social network size on performance holds even after controlling for cognitive abilities. Additionally, the results indicated that the effect of social network size cannot be reduced to amount of input, as those with larger networks did not report having more hours of verbal interaction, the amount of verbal interaction did not predict performance on the task, and the effect of network size held even after controlling for amount of input. To conclude, that study showed that similarly to the way that phonological acquisition is facilitated by exposure to multiple speakers, having a larger social network can boost phonological skills even in adult native speakers, at least when it comes to understanding speech in noise.

Another study tested whether the beneficial effect of increased social network size is limited to the phonological level, or also extends to the semantic level (Lev-Ari 2015a). Research on phonological acquisition argued that exposure to multiple speakers boosts learning because the phonological level is characterized by many-to-many mapping (e.g. Pisoni and Lively 1995). That is, two different people can produce different sounds that refer to the same phoneme, while the same sound could refer to different phonemes in different contexts. Evaluative adjectives are similarly characterized by many-to-many mapping. Two people can use the word great to mean quite different levels of enthusiasm, and use different terms differing in extremity to express the same level of enthusiasm. In other words, one person’s great can be another’s ok. Therefore, the same way in which exposure to multiple speakers assists in overcoming the difficulty of disambiguating sounds that lack one-to-one mapping to phonemes, such exposure could assist in disambiguating evaluative adjectives, which lack a one-to-one mapping to a meaning.

Results show that, indeed, having a larger social network leads to better performance at the semantic level. In particular, when reading restaurant reviews, people with larger social networks were better at estimating how many stars a reviewer assigned the restaurant, indicating better comprehension of the evaluative language of the review. Moreover, this effect was shown to be causal, as an experimental manipulation of social network size produced the same results. Specifically, participants were exposed to novel evaluative adjectives modeled after the real use of the adjectives horrible, bad, ok, good and great. All participants were trained with the same number of product reviews consisting of these novel adjectives, yet those whose reviews came from eight reviewers were better at later understanding novel reviews containing these novel words from a novel reviewer than those who were trained on reviews that were all written by only two reviewers (Lev-Ari 2015a).

So how does input from more speakers help individuals overcome the many-to-many mapping? The literature on phonological acquisition in infants and L2 learners ascribes the beneficial contribution of having multiple speakers to variability. For example, it has been argued that variability helps the learner discover which aspects of the input are relevant and which ones are not (Rost and McMurray 2010). For instance, by hearing the same word refer to the same object despite differences in pitch or intonation contours across tokens, the learner learns that pitch and intonation contour are irrelevant when mapping words to their referents. Similar arguments have been raised in the vision literature about the acquisition of visual categories (Posner and Keele 1968). In contrast, Sumner (2011) argued that input variability helps because it increases the odds that an incoming stimulus matches a stored instance. Another option is that variable input might also be more representative of the population at large, and thus facilitate the ability to deal with incoming input from novel speakers. These proposals are not a-priori mutually exclusive, though some of the evidence is conflictory. For example, Rost and McMurray (2010) found that only variation along
the irrelevant dimensions of the input (e.g. pitch, intonation) boosts learning while variation along the relevant dimension (VOT) is not useful. In contrast, Sumner (2011) found that variation along the relevant dimension (VOT) facilitated adaptation. One way to reconcile these findings is to note that Rost and McMurray (2010) focus on the initial stages of language acquisition. Their account focuses on the process of discovering the categories and patterns in the language that is learned, and their participants were, accordingly, infants. In contrast, Sumner’s (2011) account focuses more on language tuning after it has already been acquired. Therefore, her experiment included adult native speakers of the language. It might be the case then that variability plays a different role at different stages of learning.

One computational study investigated which of these proposals, if any, accounts for the benefit of having larger social networks on performance at the phonological level (Lev-Ari 2016). In that study, an agent interacted with a random subset of 20 or 100 speakers randomly sampled out of a population of a thousand speakers, who all produced Dutch vowels, represented by their formant frequencies. During the interactions, the agent stored all the productions of its interacting partners. Following the interaction phase, the agent was tested on recognizing a vowel from a speaker outside its network. The simulations replicated the experimental behavioral results, indicating that agents with networks of 100 are significantly more accurate than agents with networks of 20 at recognizing vowels. These simulations also showed that, as in the experiment, the positive effect of social network size is independent of amount of input and occurs regardless of how much interactions the agents have before they are tested. In contrast, the simulations revealed that the effect of social network size does depend on the ratio of intra- to inter-individual variability, such that having a larger social network is most beneficial when speakers are consistent within themselves, but the population is heterogeneous.

Surprisingly, the simulations revealed that the beneficial effect of network size cannot be reduced to variability, at least not when defined in terms of the variance in the input or in terms of the similarity of the members in the network to one another. When either type of variability was tested, its effect was either null or negative, depending on other features of the simulations. Instead, the beneficial effect of having a larger social network turned out to be due to a different distributional property, smooth sampling, that is, the degree to which the sampled input covers more of the central regions of the distribution. In the simulations, agents with smaller social networks received input whose distribution was characterized by scattered clusters with unsampled areas in-between them. In contrast, agents with larger social networks received input whose distribution was spread over the entire sound space, with fewer and smaller unsampled patches (see Figure 1). This distributional consequence of the network size fully accounted for the beneficial effect of social network size.

Taken together, evidence so far indicates that the more that the input we receive is representative of the language of the community at large, the better our linguistic performance will be. The less the distributional properties of the input we receive diverge from typical distribution, and the more our input includes samples of all patterns, the better the comprehension of others at all linguistics levels will be.

2.2 The role of input in modulating attention patterns and cue weighting

The structure of the input we receive influences not only our knowledgebase, but also our consequent processing manner. The cues that we attend to are those that had been predictive in our past input. If our past input matches the structure of the language of the community at large, we would implicitly learn to attend to the relevant dimensions. Otherwise, we might not realize that some dimensions are predictive, or even attend to irrelevant dimensions.

So far, the role of input in shaping our attention has mostly been investigated in the cross-linguistic literature. While not focusing on the role of the social environment or its input per se, such research has been informative in showing that exposure to different languages leads speakers to attend to different aspects of the input, because of differences in the degree that different cues predict meaning across languages. For example, languages differ in the degree to which different elements, such as word order,
animacy, or agreement, are predictive of interpretation. Consequently, speakers of different languages weigh these cues differently during processing (MacWhinney et al. 1984). Furthermore, if the structure of the input changes throughout our lives, we will accordingly adjust the weights that we give different cues or different interpretations. Such a change in manner of processing was exhibited in Spanish-English bilinguals’ attachment preferences. In particular, whereas monolingual Spanish speakers attach the relative clause in sentences, such as, *The police arrested the sister of the servant who had been ill for a while* to sister, Spanish-English bilinguals with immersion exposure in an English environment attach it to servant, as monolingual English speakers do, even when reading in Spanish (Dussias and Sagarra 2007).

The structure of the input can also influence the success of learning a property of the language. Gomez (2002) showed that input variability determines the likelihood of picking up patterns in the input. Specifically, both 18-months-old infants and adults are better at learning non-adjacent dependencies when the intervening element in the input varies a lot than when it only has few variations. This benefit of variability is argued to be due to the fact that variability prevents the memorization and storing of each element, forcing the need to generalize over items, and thus increases the likelihood of noticing patterns in the input (Gomez 2002).

To sum up, the structure of the input can shape which patterns we are likely to note and extract, and how we will weigh different cues.

**Figure 1:** Illustration of typical input that agents with a network size of 20 (*top row*) and a network size of 100 (*bottom row*) receive. The 3-dimensional input (f1, f2, f3) is broken down to an f1 by f2 vowel map (*left column*) and an f2 by f3 vowel map (*right column*).
2.3 The role of input in shaping expectations

A perhaps more surprising way in which input can influence our linguistic performance is by shaping our expectations. Linguistic patterns are structured, such that speakers’ use is predicted by their indexical features, the linguistic context and so forth. For example, women and men have different vocal tract sizes on average, leading their phonemic boundary between neighboring vowels, such as /ʊ/ and /ʌ/, to occur at different formant frequencies. Listeners rely on this difference, and apply different boundaries when interpreting the speech of women and men. Thus, when hearing tokens on the /hʊd/-/hʌd/ continuum, listeners shift their interpretation from *hood* to *hud* at lower formant frequencies when listening to male than to female voices. More importantly, listeners do so even when merely told to imagine that the speaker is a man versus a woman (Johnson et al. 1999). Similar effects of listeners’ expectations regarding the identity of the speaker on their interpretation of the speaker’s speech have been found with expectations regarding other indexical information, such as speaker’s age, race, or even occupation (Drager 2005, 2011; Koops et al. 2008; Pexaman and Olineck 2002; Staum Casasanto 2008).

While in general reliance on such indexical information is useful since production varies consistently with dimensions of speakers’ identity, listeners might also rely on stereotypes and expectations that are inaccurate, leading them to misperceive the speech. For example, in a vowel-matching task, listeners selected different tokens as being most similar to the one produced by the speaker depending on whether they believed the speaker to be from Detroit or Canada (Niedzielski 1999). More strikingly, listeners found the speech of native English speakers to be more foreign-accented (Babel and Russell 2015; Rubin 1992) and less intelligible (Babel & Russell) when it was presented with a picture of an Asian versus a Caucasian speaker.

Expectations can influence not only the way language is interpreted but even the manner in which it is processed. For instance, non-native Dutch speakers often make gender agreement or inflection errors in Dutch. Consequently, when Dutch listeners hear a native Turkish speaker compared with a native Dutch speaker make a grammatical error typical for non-native speakers, the amplitude of the ERP P600 component, which is sensitive to grammatical violations, is reduced (Hanulíková et al. 2012). There is no similar reduction in response to semantic violations that non-native speakers do not make. Other studies show that it is not only that certain patterns are expected and accepted, but that the entire manner of processing is modulated. In particular, listeners’ expectations of non-native speakers to have low linguistic proficiency can lead them to process the linguistic content, which is less reliable with non-native speakers, in less detail, and instead, increase their reliance on non-linguistic cues, which are equally reliable with native and non-native speakers (Lev-Ari 2015b). Even talker normalization is influenced by whether listeners believe that they are listening to a single speaker or two speakers (Magnuson and Nusbaum 2007).

Naturally, people may differ in their experience, and this can give rise to different expectations, and consequently, differences in processing. For example, both native German and native Dutch speakers often substitute the English /θ/ with other phonemes, but whereas for German speakers, the most frequent substitution is /s/, for Dutch speakers, it is /t/. Correspondingly, native German speakers are quicker to fixate on a target when hearing an English /θ/ substituted with an /s/ than with a /t/, while native Dutch speakers show the opposite patterns (Hanulíková and Weber 2012). Experience can also differ even within the same language community. Drager (2011) found that age expectations influence vowel perception in older but not younger participants, presumably because older adults had greater experience and awareness of the merger in progress in the community than the younger adults. Perhaps the most direct link between individual differences in experience and performance can be seen in a study that examined individual differences in people’s community make-up (Lev-Ari & Peperkamp in press). That study examined census data about the proportion of foreign speakers of two types in participants’ communities: speakers of Spanish (which has a rolled ‘r’), and speakers of other foreign language (that do not have a rolled ‘r’). The study then matched these statistics to participants’ performance on tasks measuring expectations.
regarding how likely an unknown European language is to have a non-rolled versus a rolled ‘r’, and misperception of non-rolled ‘r’ as rolled ones in a foreign language. The study showed that community make-up influences both expectations and perception. Interestingly, community make-up was shown to influence expectations not only by exposure to relevant patterns, but also by providing salient irrelevant information that can shape expectations. In particular, most non-Spanish foreign language speakers in people’s community were of Asian origin, yet this proportion influenced participants’ perception of an unknown language they knew to be European.

The patterns in our input, then, can shape our expectations, and thus later influence the way we approach, perceive and interpret incoming language. For the most part, this reliance on expectations that are derived from past input is useful, as it can allow us overcome the problem of many-to-many mapping at certain linguistic levels, and facilitate processing in general. At the same time, if the input is unrepresentative, or if the expectations are applied inappropriately, such reliance on expectations can bias interpretation and reduce accuracy.

3 General discussion

In this paper I argued that studying individual differences in people’s social environment, and in particular, in the input that the environment provides, is important for understanding individual differences in linguistic performance, as well as for understanding the mechanisms underlying language learning and language processing in general. Specifically, I delineated three ways in which input can shape linguistic performance: by influencing our knowledgebase, by modulating our attentional patterns and manner of processing, and by shaping our expectations. I propose that by paying closer attention to individual differences in people’s input, we will understand better which information we extract from the input, how we store and represent it, and how we use it when interpreting incoming language. In other words, studying individual differences in input will allow us to better understand how we learn, process and represent language.

Throughout this paper, I examined linguistic performance in terms of success at understanding unfamiliar speakers from the community or general patterns in the language of the community at large. It is likely that the relation between different aspects of the input and performance would be different when it comes to understanding familiar speakers. Therefore, depending on one’s communicative needs, having input with a certain structure would be more or less useful. Similarly, the relation between different aspects of the input and performance might be different across different societal structures, as these structures might influence the range and frequency of different communicative situations, and consequently, communicative needs. In all communities and communicative situation though, understanding the input that the social environment provides is important for understanding how linguistic skills develop and how linguistic performance is shaped.

References


