

Phonetic convergence, language talent, personality and attention

Natalie Lewandowski, Matthias Jilka

Angaben zur Veröffentlichung / Publication details:

Lewandowski, Natalie, and Matthias Jilka. 2019. "Phonetic convergence, language talent, personality and attention." *Frontiers in Communication* 4: 18.
<https://doi.org/10.3389/fcomm.2019.00018>.

Nutzungsbedingungen / Terms of use:

CC BY 4.0





Phonetic Convergence, Language Talent, Personality and Attention

Natalie Lewandowski^{1*} and Matthias Jilka²

¹ Experimental Phonetics Group, Institute for Natural Language Processing, University of Stuttgart, Stuttgart, Germany,

² Applied English Linguistics, Department of English and American Studies, University of Augsburg, Augsburg, Germany

OPEN ACCESS

Edited by:

Pia Knoeferle,
Humboldt University of Berlin,
Germany

Reviewed by:

Zofia Malisz,
Royal Institute of Technology, Sweden
Caicai Zhang,
Hong Kong Polytechnic University,
Hong Kong

*Correspondence:

Natalie Lewandowski
natalie.lewandowski@
ims.uni-stuttgart.de

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 30 April 2018

Accepted: 24 April 2019

Published: 15 May 2019

Citation:

Lewandowski N and Jilka M (2019)
Phonetic Convergence, Language
Talent, Personality and Attention.
Front. Commun. 4:18.
doi: 10.3389/fcomm.2019.00018

Studies into phonetic adaptation rarely consider individual differences (IDs) on a cognitive and personality level between speakers as a direct source of adaptation variation. In order to investigate the degree to which the individual phonetic talent and further psycho-cognitive IDs of speakers affect phonetic convergence in a second language setting, 20 German native speakers were involved in two dialog tasks with two native speakers of English, a male speaker of American English and a female speaker of Standard Southern British English. The dialogs were quasi-spontaneous task-oriented interactions elicited with the Diapix picture-matching game. The English L2 learners were divided into a phonetically talented and less talented group based on their test results and evaluation in a preceding extensive language talent test battery. The acoustic analyses using amplitude envelopes revealed that talented speakers converged significantly more toward their English native speaking partners in the Diapix study. An additional analysis relates their degree of convergence to a range of personality and cognitive measures. The factors openness, neuroticism, Behavior Inhibition score and the switch costs in a Simon Test significantly impacted the degree of phonetic convergence in the dialog study.

Keywords: phonetic convergence, accommodation, phonetic talent, personality, cognition, inhibition, openness to experience

INTRODUCTION

Adult language learners vary greatly in the quality and speed of acquiring the sound system of a second language (i.e., all segmental and prosodic manifestations on both the phonetic and phonological level). There are numerous possible causes for these differences. The “classic” studies of Foreign Accent have focused on external circumstances of learning such as age of learning, age of arrival, length of residence or amount of L1 and L2 use, but of course other aspects that address learners’ individual characteristics and abilities (e.g., intelligence, personality factors such as extraversion or empathy, attitude or motivation) are also well-researched. The clearest manifestation of such individual learner characteristics is of course the assumption of language or rather pronunciation learning abilities inherent to the speaker, i.e., phonetic talent or aptitude. The successful acquisition of an L2 pronunciation requires, on the one hand, the ability to correctly perceive the phonetic characteristics of that L2, and, on the other hand, the ability to faithfully reproduce these characteristics in her or his own speech. The described perception-production loop is also what characterizes the phenomenon of phonetic convergence, or phonetic adaptation, within a conversational situation—where two talkers become more alike in their pronunciation in the course of a dialog (Pardo, 2006). It may thus be the case that speakers being especially good at converging to their speaking partner during a conversation, also become very good at acquiring

the pronunciation of a new dialect or language. Or, in other terms, speakers already displaying a near-native accent in a second language, might be very good phonetic convergers, when observed, for instance in a native-nonnative dialog situation.

Phonetic Convergence

Phonetic convergence describes a process in which the pronunciation of directly interacting partners becomes more similar to each other. Acoustic measurements have shown speakers to converge on a range of global prosodic and finer-grained segmental features, including utterance duration (Matarazzo et al., 1963; Cappella and Planalp, 1981), response latencies (Street, 1984), pause duration, speech amplitude and turn-taking (Natale, 1975a,b), articulation rate (Schweitzer and Lewandowski, 2013), long-term average spectra (LTAS; Gregory, 1983, 1986; Gregory and Webster, 1996), voice onset time (VOT) and the amount of voicing in vowels (Nielsen, 2011), vowel formant values (Delvaux and Soquet, 2007; Babel, 2012; Schweitzer and Lewandowski, 2014), vowel duration and MFCCs (Delvaux and Soquet, 2007) as well as more global measurements of spectral properties, as amplitude envelope signals (Lewandowski, 2012).

The first wave of studies on speaker adaptation in the 1970s arose from Speech Accommodation Theory (later on *Communication Accommodation Theory*, e.g., Giles and Powesland, 1975; Simard et al., 1976; Giles, 2001), where accommodation toward a speaking partner was seen as a way to create a socially comforting environment, i.e., boosting social attractiveness and reducing social distance (Giles, 2001). CAT distinguishes between positive accommodation (convergence) as a means of reducing social distance, divergence—a means increasing social distance, and maintenance—keeping one's own style (Giles and Ogay, 2006). Convergence also expresses the need for social approval, group membership and can enhance communicative effectiveness (e.g., by agreeing on joint vocabulary; Pitts and Giles, 2008). Although the tenets of CAT might imply a certain amount of control over the process, it has not yet been established how far this control extends and which types of linguistic changes might be more prone to the influence of social and psychological factors than others (Lewandowski, 2012).

Pickering and Garrod (e.g., 2004, p. 2, 2004, p. 20, 2005, 2006, 2013) see alignment as largely automatic and very straightforward, whereas other influences (i.e., social or personality-related) are not discussed. Their mechanistic model proposes that comprehension and production—or in later extensions also listener and speaker expectations and predictions (*forward model*; Pickering and Garrod, 2013)—become coupled or synchronized during a conversation, therefore a.o. facilitating mutual comprehension (Pickering and Garrod, 2004). This account is in line with studies on (social) coordination dynamics which also foresees biological hard-wiring of general coordinating behavior in humans (e.g., Kelso, 1997; Kelso and Engström, 2006). Yet another strand of research into convergence—the hybrid accounts—do not deny a biological core mechanism but concurrently allow for social

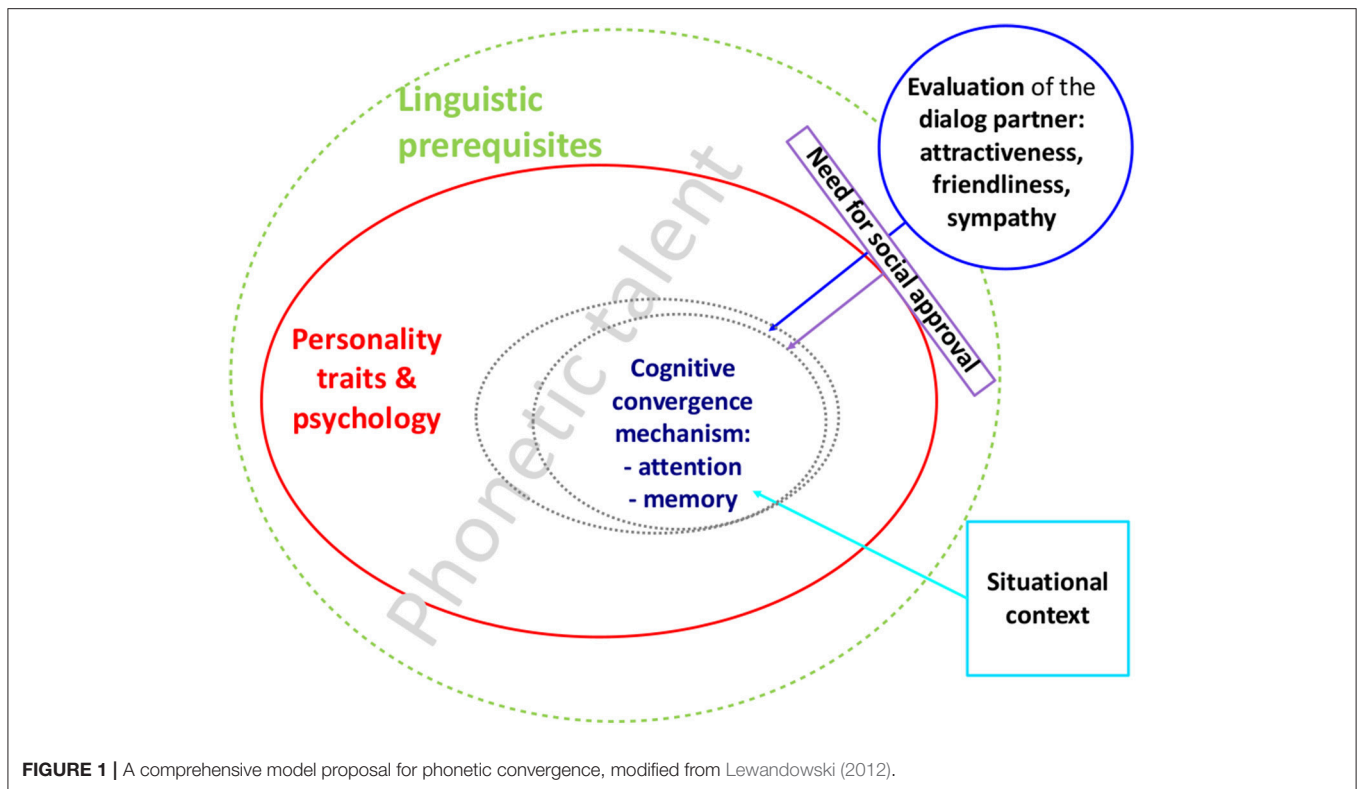
and psychological influences (e.g., Krauss and Pardo, 2004; Lewandowski, 2012; see also **Figure 1**).

The elicitation techniques, or scenarios, in which accommodation processes were measured, also varied over the last 50 years. As accommodation in the CAT terminology involved the presence of a social situation, the effects of adaptation in the early studies were also investigated almost exclusively within conversational interactions: within *free* or *semi-free conversations* (Cappella and Planalp, 1981; Gregory, 1983, 1986), with a hidden experimenter (Natale, 1975a), and *interviews*, in either a laboratory (e.g. Natale, 1975a,b; Street, 1984; Willemyns et al., 1997) or a quasi-natural setting, as e.g., Larry King's talk show guests (Gregory and Webster, 1996). In more recent times though, a considerable amount of studies on phonetic speaker adaptation was, for instance, based on shadowing paradigms (see Pardo et al., 2017 for a recent review of non-interactive designs). These experimental set-ups involve the repetition of words or longer stretches of speech (Namy et al., 2002; Delvaux and Soquet, 2007; Nielsen, 2007, 2008; Babel, 2009; Brouwer et al., 2010; Abrego-Collier et al., 2011), which are either based on modified versions of Goldinger's shadowing paradigm (Goldinger, 1998), or, for instance, on short question-answer sequences [word games, as e.g., the dominoes game in Bailly and Lelong (2010)]. Another type of set-ups renders dialogs of a still limited linguistic nature (both in terms of utterance length and complexity) and works with assigned talker roles: Map Tasks (Pardo, 2006; Smith, 2007; Pardo et al., 2012, 2013, 2018; Aguilar et al., 2016) and interactive search tasks and games (Dias and Rosenblum, 2011; Levitan et al., 2015). The last group of studies relies on quasi-spontaneous¹ or fully spontaneous conversational data (De Looze et al., 2011; Kim et al., 2011; Schweitzer and Lewandowski, 2013, 2014; Schweitzer et al., 2015).

Pardo et al. (2018) provide an argument for indeed turning toward more natural data for research on pronunciation accommodation by concluding that imitation effects for shadowing² vs. convergence effects in more interactive forms of communication (MapTasks in this case) might in fact not be directly comparable. Phonetic convergence arising in fully natural contexts—in interactive, dynamically evolving dialogs—could thus be even further away from the effects we find in imitation studies, e.g., in shadowing. Taking into account the greatly distinct demands these two scenarios pose for a listener-speakers' attentional system, which is crucial for the storage and retrieval of appropriate exemplars (be it words, sounds or phrases; see section Individual Cognitive Differences in an Exemplar-Theoretic Account of Language), we would similarly like to encourage a differentiation between the terms *imitation*, as the measurable adaptation arising in highly linguistically controlled and scripted environments, with limited speech material at hand, and *phonetic convergence*

¹Kim et al. (2011) use Diapix tasks, which will be further described in the methodology section.

²The direct (or delayed) imitation of (usually) short phrases, words, syllables or sounds, in a task-based manner and not following from natural, unscripted conversation.



arising in natural interactions, while juggling the complexity of meaningful conversation with a dialog partner. We assume that an excellent performance in a shadowing experiment does not automatically indicate or predict an equally high degree of phonetic convergence in running unscripted speech, where attentional resources are necessarily divided upon considerably more parallel tasks, are largely directed at (or away from) pronunciation features naturally (i.e., implicitly), and a number of additional social and contextual factors might enter the equation (see also Lewandowski and Duran, 2018). Throughout this paper, convergence will thus be referred to meaning spontaneous, naturally emerging adaptation in complex *conversational interaction*.

Factors Influencing Phonetic Convergence

A considerable amount of factors has been studied in conjunction with phonetic adaptation, predominantly those of an interactive nature, i.e., where the dialog partner or the social context move into the foreground (see schematic model in **Figure 1**). Amongst such well-investigated social factors resulting in higher or lower degrees of convergence, are for instance dialog partner evaluations (e.g., friendliness, attractiveness; Pardo et al., 2012; Schweitzer and Lewandowski, 2013, 2014) or social preferences (e.g., racial or sexual bias; Babel, 2009, 2012; Abrego-Collier et al., 2011). In stark contrast to this lies the almost non-existent body of literature dealing with psychological and personality characteristics of the speaker, or his or her cognitive individual differences and their role in convergence, although they form an essential part of a full model of phonetic convergence. Lewandowski (2012) proposed a model of phonetic convergence

including not only social and contextual factors but also linguistic skills, psychological features and a cognitive core mechanism, which might be related to attention and working memory (WM) skills (see **Figure 1**).

So far only one study (Yu et al., 2013) reports a correlation between the amount of imitation and the personality feature “openness,” and a further correlation to a test variable from the Autism Spectrum Quotient (AQ, Baron-Cohen et al., 2001) which can be related to attention-switching. Yu and colleagues’ study, however, does not look into conversational interaction but is an exposure study. The measure associated with attention-switching is based on a subjective self-questionnaire and not an objective and test-based metric. Yet no study in phonetic convergence in conversational interaction to date has been devoted to the psychological and cognitive individual differences (IDs) of the speaking partners, which is the main focus of this paper.

Individual Cognitive Differences and Phonological/Phonetic Processing

Although not directly within speaker adaptation studies, the impact of varying attention skills has been investigated in a number of studies on speech perception and production, encompassing clinical and normal populations, and L1 as well as L2 speakers and in plurilingual contexts.

The relationship between bilingualism and attention, for instance, has been put forward in Green’s model of Inhibitory Control (Green, 1998) where bilinguals are said to be permanently faced with the need to inhibit the currently not used language to be able to communicate in the other—a

strategy subject to individual differences (Gollan et al., 2011; Lev-Ari and Peperkamp, 2013). Green's model also foresees asymmetric switch costs, an effect which can be modulated by the individual's better inhibition skill, too: speakers with higher inhibitory control scores obtain lower switch costs when shifting between languages (Linck et al., 2012). Looked at from another perspective, bilingual speakers do seem to profit from a cognitive processing benefit, evident from generally lower switch costs in a Simon Test and faster RTs in tasks with high working memory demands (Bialystok et al., 2004). The impact of differing attention skills in L2 and bilingual contexts was also studied for VOT productions (Lev-Ari and Peperkamp, 2013), the perception and production of vowels and consonants (Darcy et al., 2016; Safronova, 2016), and tone perception and production (Ou et al., 2015; Ou and Law, 2017). The effects of poorer attention (switching or inhibitory) control bearing on speech processing have been also previously documented in elderly listeners, leading to stronger perceptual learning effects (Scharenborg et al., 2015). Speech development studies have furthermore found evidence for the essential role attention plays in the fine tuning of perceptual representations (Jusczyk, 2002; Conboy et al., 2008). Analyses of attentional influences in clinical populations, as in Specific Language Impairment (SLI) or developmental dyslexia, brought forward the Sluggish Attentional Shifting Hypothesis (Hari and Renvall, 2001; Lallier et al., 2010) which suggests that dyslectic patients suffer from a prolonged cognitive integration window that induces difficulties in decoding the correct temporal sequence of (speech) units and leads to disturbed rapid stimulus sequence processing (RSS).

Individual Cognitive Differences in an Exemplar-Theoretic Account of Language

Individual differences in cognition and personality leading to differential linguistic performance, including phonetic convergence, can be very straightforwardly modeled in an exemplar-theoretic account—falling within usage-based accounts of language. Exemplar Theory incorporates recency effects, which are so crucial within speaker adaptation, and can easily aid to explain the straightforward link between perceiving and reproducing someone else's utterance within a dialog (Goldinger, 1996, 1998; Johnson, 1997; Hawkins, 2003; Pierrehumbert, 2006). The stored exemplars of speech include rich indexing with, amongst others, labels for speaker identity, language, dialect, accent or situation. Within such a framework the dialog partner has access to an immense pool of exemplars to choose from—in order to adapt to their partner (by finding a closely matching exemplar) or not to adapt (by not finding or not choosing a closely matching exemplar). Furthermore, it is already at the acquisition stage of an exemplar, where many intermediate steps come into play, introducing much room for IDs between listeners and their prospective phonetic adaptation behavior (Pierrehumbert, 2006). Lastly, the adaptation process happening between two dialog partners form the microscale, or, the starting point of any further-reaching macroscale processes, as language change – which have been already modeled in usage-based accounts (e.g., Bybee, 2002, 2006).

Within such a usage-based account of language, Pierrehumbert (2006) draws our attention to a suitable solution for the “perfect imitation” problem (i.e., no two instances of the same utterance, word or sound are ever pronounced identically). Her proposal further divides the acquisition stage of exemplars into intermediate steps of *noticing*, *recognition* and *coding*, forming the basis for the multiple mechanisms standing in between the mere physical experience of a stimulus and its subsequent re-usage in production (Pierrehumbert, 2006). The incoming exemplars are first subject to the operation of *noticing*. It is the first and crucial processing step on which all subsequent steps rely. So it seems we have identified an especially powerful cognitive mechanism when it comes to explaining possible IDs in exemplar processing and phonetic convergence—namely attention. Hawkins (2003) argues that every person (listener, speaker) develops a distinct mental representation of language, simply because categories are self-organizing and emergent. It follows that every individual is faced with a varying linguistic (and also phonetic) input, and not every person seems to have equal control over the attentional and working memory mechanisms (Robinson, 2003; Cowan et al., 2005; Styles, 2006) which does inevitably surface in processing difficulties of fine acoustic detail necessary for acquiring a native-like pronunciation. Vais et al. (2015), for instance, found a link between speaker talent and accurate use of frequency information in the L2 in their variability study. This implies that talented L2 learners are better able to build categories from exemplars in the L2. Considering the output—or—retrieval stage of exemplars, Skehan (2003) emphasizes a further advantage of an exemplar-based route over a purely rule-based access system in natural conversations: although the exemplar route might be less flexible and rely on chunks and redundant storage, it is fast and provides convenient access, forming, in his opinion, the basis for both native-like selection and fluency (Skehan, 2003).

Summarizing the links between exemplar-theoretic models and convergence mechanisms, we propose that speakers might be differently endowed with certain cognitive skills (e.g., attention) enhancing or hampering the correct noticing of richly indexed exemplars, as well as their subsequent storage and retrieval in running conversation, leading (amongst other crucial factors, as the aforementioned psychological features) to differences in their degree of phonetic convergence. Lewandowski (2012) proposes that insufficient attention on fine phonetic detail and the focus shifted toward meaning recognition could interfere with the proper recognition and storage of new acoustic-phonetic exemplars, leading to incomplete or poorly indexed exemplars. As a consequence, no (sufficient) top-down modulation³ would be possible to help in identifying the currently best items (exemplars) for production at the output stage, since the acoustic-phonetic information (on speaker, accent, dialect etc.) would simply not be stored. This suggests another possible difference in the processing of acoustic-phonetic information by talented and less talented language users.

³Top-down modulation used in the meaning adopted by Adaptive Resonance Theory (ART, e.g., Grossberg, 2003, see Figures 5, 6).

The Role of Phonetic Talent

The special status of phonetics in L2 acquisition has led to the assumption of a distinct talent component responsible for a person's success in the L2 phonetics. The same talent factor might also be involved in the mechanism controlling phonetic convergence in a speaker.

Within this concept of language talent L2 research widely accepts the special status of phonetic talent as opposed to other linguistic abilities. Typically a fundamental distinction is drawn between two main substrates of linguistic ability: talent for grammar vs. talent for accent (Schneiderman and Desmarais, 1988). The additional difficulty of pronunciation acquisition in contrast to other linguistic features is well-known as the so-called "Joseph Conrad Phenomenon" (e.g., Guiora, 1990; Bongaerts et al., 1995; Abu-Rabia and Kehat, 2004), which refers to the Polish-born novelist's native-like abilities in English grammar (syntax, morphology) and vocabulary apparently having being accompanied by a strongly accented pronunciation. This separation of pronunciation from other L2 skills has been confirmed in a number of experimental studies. Neufeld (1987), for example, showed that ratings of pronunciation skills did not correlate with the results of general language aptitude tests. The phonetic subsystem is generally thought to be more difficult to acquire, as it is assumed to rely more on hard-wired biological processes that cannot easily be influenced by conscious learning efforts. Also, in the acquisition of L2 pronunciation there is the additional challenge of having to bypass already established sensory and motor pathways in order to either correctly perceive speech sounds or control articulatory movements.

It is also generally agreed that a distinction has to be drawn between proficiency, i.e., the overtly observable performance of a particular skill, and innate talent or aptitude. Given the same learning circumstances and similar language experience backgrounds, some learners will inevitably be better than others. Talent or aptitude is thus defined as a stable, innate characteristic that is separate from external circumstances of learning such as experience, input, age of learning etc. as well as other attitudinal factors and personal abilities such as motivation or intelligence. While it is not an indispensable prerequisite for SLA, it does crucially enhance rate and ease of learning (Carroll, 1981). According to Dörnyei (2005) talent manifests itself via an ideal (or at least very effective) processing of learning conditions and novel information in terms of higher-order cognitive processes (like analysis and inference), lower-order cognitive processes (like pattern recognition) and specifically phonetic abilities like hearing, perception, articulatory flexibility and memory of sound features. As stated earlier, there is, however, no comprehensive model explaining phonological talent, as it is unclear how phonological skills interact and develop over time (Moyer, 1999). Proficiency, on the other hand, can be ascertained more directly as the sum of both inherent (such as, for example, talent), and external influence factors (e.g., amount of L1 and L2 use) in an overall performance test (which of course would not reveal the interactions between all these factors).

In order to control for the effect of experience-related factors, it would appear to be the best possible course of action to assemble a large group of test subjects that is homogenous at least with respect to age and "learning career", i.e., identical time and circumstances of the acquisition of the L2 and then collect as much detailed information as possible on these and all other potentially influential factors such that any correspondences with performance would not remain undetected. This would also include tests for cognitive (e.g., working memory, intelligence etc.) and socio-psychological (e.g., personality traits) parameters (see the following section on assessment for details), and a measure for motivation.

There are, of course, also tests that are designed to measure general language ability directly or rather predict success at learning a second language, mainly on the basis of the L1 (as the Modern Language Aptitude Test, short MLAT, Carroll and Sapon, 1959). With regard to a direct assessment of phonetic talent, on the other hand, such tests or a combination of tests have yet to be established. In fact, one of the objectives the extensive investigation of the neural correlates of pronunciation talent (Dogil and Reiterer, 2009) referred to in later sections had, was to gain a better insight into specific abilities or tasks that would be particularly representative with respect to that purpose. It is also in this sense that the ability of some speakers to adapt to the phonetic characteristics of other speakers (as defined in Subjects) might be interpreted as an expression of talent, while it (i.e., phonetic convergence) might nevertheless in turn be connected to and/or influenced by yet other individual factors, as well as contextual and social factors (see section Factors Influencing Phonetic Convergence).

Motivation and Goals

As described in the introductory part, the impact of social factors on phonetic adaptation (including but not limited to social and professional status, mutual liking, and aspects of dominance) has so far received considerably more attention within studies on phonetic adaptation than personality features and cognitive factors—including pronunciation talent. The phonetic adaptation mechanism though, lying at the intersection of perception and production, is in our opinion susceptible not only to social and contextual influences (external factors) but also, and maybe foremost, to factors lying within the listeners-speakers themselves (internal factors). Analyzing the convergence mechanism from an exemplar-theoretic angle, all steps of the perception-production loop (noticing, recognition, (en)coding, retrieval and production) pose great demands on the language user's cognitive functions, which differ largely from one person to another. We assume that individual differences in executive attention, and possibly also working memory components, have a considerable impact on the degree of convergence. The skillful and fast directing of attention toward the relevant acoustic-phonetic features is crucial for success in both storage and retrieval of exemplars. Such a competent employment of attentional resources could allow for the storage of richer-indexed exemplars, retaining more relevant acoustic-phonetic features, which can later on be accessed for production

(on top of facilitating the retrieval process itself). Precisely this skill—a better employment of executive attention—might also be the one setting apart the phonetically talented speakers in our study from the phonetically less talented ones, leading to potentially more phonetic convergence in the former group. A person striving for (be it consciously or subconsciously) for convergence, must first possess the necessary skill to do so—a certain aptitude for fast storage and re-usage of acoustic-phonetic material. Observed the other way round: a good converger, adapting to their speaking partner's pronunciation in an L2, has a high chance of being a talented and successful acquirer of this L2's proper pronunciation in general. In addition to that, we assume that some personality features can further facilitate this convergence process, by potentially mediating the top-down directed attention at the partner's language, and more specifically at her or his pronunciation. This, again, influences the amount of adaptation possible—or desired. We would, for example, expect personality features as *openness* (see also Yu et al., 2013) and *extraversion* to positively impact the degree of convergence, and measures as the *Behavior Inhibition Scale* (BIS) to stand in a negative relationship to the amount of convergence.

In what follows, we will first describe the procedures of the background study (also referred to as *the language talent study* from this point onwards) on *Language Talent and the Brain* (see Dogil and Reiterer, 2009). Our current study on phonetic convergence and the language talent study share participants. The subjects for the convergence study have been chosen amongst the large pool of participants of the talent study, guided by their performance therein. Apart from the linguistic tests allowing for a classification of the speakers into a talented and less talented group, the language talent study also contained a number of additional psychological and cognitive data on the participants, in parts selected for the analysis of convergence in the main study. The second part will be concerned with the methodology of our main study—the analysis of phonetic convergence in relation to the phonetic talent of our participants, as determined through the linguistic tests in the language talent study (Dogil and Reiterer, 2009; Jilka, 2009a), and to their personality and cognitive skills.

LANGUAGE TALENT STUDY: IDENTIFICATION AND CLASSIFICATION OF PRONUNCIATION TALENT IN LANGUAGE LEARNERS (Dogil and Reiterer, 2009)

As indicated in the preceding section, the selection of suitable participants for the phonetic convergence study (our main study, described in section Main Study: Testing Phonetic Convergence in a Conversational Setting) was based on an earlier project that had the objective of providing a comprehensive examination of talent in second language pronunciation—the language talent study (Dogil and Reiterer, 2009; see also Jilka et al., 2010). It investigated this ability with respect to its multiple phonetic/linguistic manifestations, attempting to take into account the large variety of influence factors. The main

goal was to facilitate further investigations into finding neural correlates of pronunciation talent, i.e., differences in brain activity between talented and untalented speakers.

The Test Set-Up

This pronunciation talent investigation is described in detail in Jilka (2009a) and employed 102 native speakers of German as test subjects including a core group of 50 university students of English who shared a number of key variables such as age (range: 20–23 years), age of onset of L2 English learning (10 years) and type of experience/training in this L2 (English instruction in the formal setting of the German school system, relatively low amounts of experience in English-speaking environments). Further subjects from outside university were chosen based on their self-professed pronunciation talent or lack thereof. Fifteen native speakers of English also participated in order to provide data for comparisons. All subjects were adults and were informed in writing about their rights following the recommendations of ZENDAS (Zentrale Datenschutzstelle der baden-württembergischen Universitäten). All subjects then gave written and informed consent to participate in the study in accordance with the Declaration of Helsinki. The subjects received remuneration or course credit for their participation (in case of the student participants). No subjects took advantage of their right to have their measurements deleted. The study was additionally approved by the local ethics committee from the University of Tübingen, since parts of it involved neurolinguistics measurements (a comparable permission is not needed for psycholinguistic studies under German law).

The various speech tasks involved English, German, and Hindi. English was the main test language because the large number of learners increased the likelihood of finding individuals with native-like pronunciation skills. Furthermore, comparative (English vs. German) linguistic descriptions, both for the segmental and prosodic characteristics were easily available. The German and Hindi tasks were intended to restrict the influence of learning experience on performance. As native speakers of German the test subjects should have had very similar, if not equal amounts of experience with it, whereas Hindi was a language they were all completely unfamiliar with.

The complete test battery covered a wide range of conditions correlating with phonetic talent, such as the external circumstances of second language acquisition (e.g., age of learning, amount of language use, type of instruction), the learners' psychological characteristics, as well as general proficiency in the examined languages. Actual phonetic abilities were covered within the categories of production, perception and imitation (Jilka, 2009a,b).

Psychological and Cognitive Aspects

A considerable number of individual psychological characteristics have been shown to correlate with L2 performance, therefore a variety of established psychological tests was administered to the subjects in order to replicate these results and possibly relate them to each other. Two major types of tests were distinguished: those examining psychological aspects of the test subjects' personalities and those focusing on

TABLE 1 | Overview of all tests contained in the convergence study (all measurements were supplied to the statistical models as continuous data).

Acoustic measures	Amplitude envelopes to determine degree of convergence between interactants
Cognitive measures	Phonological Working Memory: digit span forward, digit span backward and non-word repetition span Attention/Mental flexibility: Simon Test (RTs)
Personality measures	Big Five: Neuroticism Extraversion Openness Five Factor Inventory (NEO-FFI) Behavior Activation/Inhibition: BIS/BAS scale Empathy: E-Skala questionnaire

cognitive abilities. A subset of these tests was chosen for the current convergence study, since they might shed more light on personality-related and cognitive IDs influencing phonetic adaptation: the *NEO-FFI*, the *BIS/BAS scale*, the *empathy questionnaire E-Skala*, *phonological WM*, and a *Simon Test* (see also **Table 1**; for a description of the original test battery see Hu and Reiterer, 2009 and Rota and Reiterer, 2009).

The *Neuroticism Extraversion Openness Five Factor Inventory (NEO-FFI)* assesses personality traits according to five major factors, namely neuroticism, openness, extraversion, agreeableness and conscientiousness (Costa and McCrae, 1992). In the current study the German (translated) version of the test was used (Borkenau and Ostendorf, 1993). The *BIS/BAS questionnaire* (Carver and White, 1994), on the other hand, investigates whether a person's actions are driven more by the positive motivation toward something desired (Behavioral Activation System-BAS) or the wish to avoid a negative experience (Behavioral Inhibition System-BIS, German adaptation by Strobel et al., 2001). A questionnaire assessing empathy, *E-Skala* (Leibetseder et al., 2001), was administered to examine if this characteristic is associated with a greater readiness to adapt to unfamiliar phonetic features, and to the conversational partner. *E-Skala* provides a general measure (*E-Skala General*) and two subscales: the *E-Skala Social* (reflecting "social concern") and the *E-Skala Empathy* (reflecting "readiness for empathy"). As far as cognitive abilities are concerned, tests of phonological working memory (Gathercole et al., 1994) were also carried out, as it has repeatedly been argued that "phonological short term memory" can predict success in L2 as well as L1 learning and phonological processing (e.g., Baddeley, 2003; Ou et al., 2015; Ou and Law, 2017; Serafini, 2017). For the current study measures for subjects' digit span forward, digit span backward and non-word repetition span were gathered. The second cognitive component—attention (inhibition)—was measured with a *Simon Test* (Simon, 1990, test for *mental flexibility*) which is a nonverbal test for *inhibition*, a cognitive component much discussed for its influence on phonological processing skills (e.g., Lev-Ari and Peperkamp, 2013; Darcy et al., 2016).

Testing Phonetic Ability

The tasks intended to assess the test participants' phonetic abilities were carried out in 90 min sessions in the anechoic

chamber of the Institute for Natural Language Processing at the University of Stuttgart (Jilka, 2009a). Various elicitation techniques were applied, following the example of studies like Oyama (1976), Bongaerts et al. (1995), Markham (1997) or Flege and Hillenbrand (1987), to elicit differing types of intonational configurations, speaking rates, and degrees of fluency. Especially imitation and reading tasks were employed to elicit tunes (i.e., combinations of pitch accents and boundary constellations) associated with particular discourse situations (e.g., declaratives, Yes/No-questions, continuation rises). Insights gained in previous work (e.g., Jilka, 2000) facilitated the identification of possible deviations due to foreign accent such as tonal category transfer or variation in the phonetic realization of equivalent categories.

Thus, the test battery (see also **Table 2**) included three major blocks, namely tasks for speech production, speech perception and imitation task (combining perception and production). Speech production tasks focused on the elicitation of (quasi-) spontaneous speech (reflecting natural fluency, speaking rate, segmental realizations and choice of prosodic patterns) as well as of read speech (in order to ensure a reasonable control of segmental production and basic pitch patterns associated with particular discourse situations). Perception tasks tested the comprehension and interpretation of suprasegmental features, while imitation tasks, which obviously require most directly the correct perception of a model and the ability to reproduce it, were carried out in native (German), well-known (English) or unknown (Hindi) language, covering both segmental and prosodic detail in direct and delayed set-ups.

The results of the many different task types manifested themselves in multiple forms. For pure production tasks involving reading, narration and acted speech, only a subjective perceptual assessment by a large number of raters was considered appropriate. For those tasks with more narrowly defined objectives, such as accent imitation or interpretative reading, it was more suitable to perform an expert analysis by referring to expected prosodic models. This applied to an even larger degree to imitation tasks where instrumental analysis allowed for the comparison with the respective originals. In perception tasks, on the other hand, evaluation was simply achieved through an automatic scoring system, as these tasks were designed to contain a fixed number of answers.

Talent vs. Proficiency

The distinction of the concepts of talent and proficiency was of course a key challenge of the analysis. The results of both those tasks designed to focus on talent and those simply measuring general proficiency in a particular area had to be related to information about the speakers' experience and degree of motivation as determined in the questionnaires and the psychological tests. While the factor of experience could be controlled to a certain extent by tasks involving either the learners' native language German, where it should be equal, or Hindi, where it was zero for all test participants, this was not possible for motivational factors. For this reason, it was concluded that there may be no single clearly-defined experimental method that directly assesses exclusively phonetic

TABLE 2 | Tests of phonetic ability [(Jilka, 2009a,b), in (Dogil and Reiterer, 2009)].

Production Tasks	(Quasi-)Spontaneous Speech (English)	Retelling of cartoon, Conversation about stays in English-speaking countries
	Reading tasks	English: "The Northwind and the sun", extract from a dialog in a short story German: "Der Nordwind und die Sonne" (imitating an English accent), short phrases (imitating English, French, Italian accents)
Perception Tasks	Prosody Pair Comparison	Tonal categories differences (English and German; also in low-pass filtered versions) Differences in the phonetic realization of identical tonal categories (English and German; also in low-pass filtered versions)
	Prosody Interpretation	Interpreting the meaning or emotion expressed in an utterance (English only)
	Accent Identification	Identifying native language of readers of "Der Nordwind und die Sonne"
Imitation Tasks	Direct Imitation	Hindi words and phrases German prosody English prosody (RP/SSB and GA)
	Delayed Imitation	German prosody English prosody (RP/SSB and GA)

talent, but that it can at best be approximated by the combination and weighting of many different tests. The variety of these tests in relation to information about the speakers' histories of language acquisition and language use as well as the aforementioned cognitive and psychological test allowed for such a weighting. It showed that tasks involving the perception of minute phonetic detail in (low-pass filtered) intonation were not helpful as they only correlated with the test subjects' ratings for the personality factor of conscientiousness. On the other hand, tasks involving general language (accent imitation and identification) and prosody (interpretation and imitation) awareness were generally good indicators of high scores in the other fields, i.e., showed the widest range of correlation.

Classification of Speakers According to Phonetic Talent

Based on the resulting perceptual assessments, instrumental analyses and automatic scoring systems (Jilka, 2009a), a ranking/classification of the speakers necessary for the subsequent neuroimaging studies of talent and proficiency was created, which eventually also facilitated the selection of speakers available for our current study. For the convergence study, 20 at this point available speakers were chosen on the basis of their phonetic talent ranking—which is composed of the summed z-scored values of the separate phonetic tests (z-scoring by test, to ensure comparability of the at times highly differing scaling systems). The participants chosen fall into two groups: phonetically gifted speakers (rated higher, mean total phonetic

score of 14.257 (z-score across all phonetic tests, $sd = 5.413$) and phonetically less talented speakers (lower end of the rating, mean total phonetic score of -9.929 (z-score across all phonetic tests, $sd = 12.856$). At first glance, two speakers toward the middle part of the spectrum performed very similarly in their total phonetic scores. For one of the speakers though, this score was largely driven by his good performance in the German parts of the phonetic test battery. For this reason, we further compared the English free speech test scores (including the introduction and the cartoon retelling) of the two subjects, which differed considerably, and used those as a further indicator for final grouping (i.e., the subject with significantly higher scores was assigned to the talented group).

MAIN STUDY: TESTING PHONETIC CONVERGENCE IN A CONVERSATIONAL SETTING

The experimental procedure of the phonetic convergence study is based in parts on the dissertation of Lewandowski (2012). An exception is the statistical analysis, which has been recalculated for the purpose of this study using linear mixed effects modeling (lmm).

The elicitation technique for the main task in the experiment needed to yield quasi-natural spontaneous speech, and still provide enough material for conducting detailed acoustic analyses at word-level. Since the conversations were between native and highly-proficient non-native speakers of English, and as this already prescribed a certain (intended) misalignment of status positions (linguistic expert vs. learner; Davies, 2003; Park, 2007), we decided for a technique which would not impose any further predefined roles. For this reason Map Tasks (Anderson et al., 2001), though previously used in convergence studies (e.g., Pardo, 2006; Smith, 2007; Pardo et al., 2012), were not considered, for they have a fixed talker role assignment of Instruction Giver and Instruction Follower, which brings about considerable disparities in the length and total amount of vocalizations. Furthermore, also the quality of the utterances, in terms of the variability of utterance types and their complexity (e.g., the frequently observed domination of one-word-answers), was judged to be unsatisfactory in Map Tasks. Therefore the Diapix method (Bradlow et al., 2007; Engen et al., 2010) was eventually chosen as the best alternative for the compromise between retaining naturalness and spontaneity of dialog and still being able to ensure the repeated usage of a large number of content words for acoustic analyses involving amplitude envelope signals (described in section Amplitude Envelope Measurements).

Materials

Diapix is a "spot-the-difference" game of the type often found in newspapers and magazines (Bradlow et al., 2007; Engen et al., 2010). Each colorful set consists of two pictures which differ from one another in ten details (items have different names, colors/shapes, are dislocated, or completely missing). For the current experiment two sets of Diapix were used—the shop scene and the farm scene. The usage of two sets was necessary because

every nonnative speaker was engaged in two separate dialogs with two different native speakers of English. Although the two English native speakers took part in all 20 conversations with the German subjects and learned the location of all target items, they were instructed not to reveal this and continue acting as if they shared the same knowledge status as the nonnative speaker. Measurements for the non-native speakers, which are not part of this paper, can be found in Lewandowski (2012).

Procedure

The pictures were provided in DIN-A4 size and were laminated to reduce any interference from noise during the recordings. The subjects were instructed to solve the task of finding the target items in English and informed that there was no time limit. The Diapix task required an intensive interaction between the two conversational partners in which both speakers were free to describe what they saw and ask the other questions at any time. The words used for measuring phonetic convergence were mostly the target words appearing within the changed/missing item spots in the pictures, as e.g., bird, house, chicken, carpet, dog, but also other content words which came to be frequently used by the speakers, which ensured a sufficient number of repetitions by both speakers during the conversation, a prerequisite for carrying out the acoustic analyses.

The recordings took place in a sound attenuated chamber, with the participants separated by a padded wall. The interactants could not see each other, but heard each other clearly through headphones. Once the dialog started, there was no interference from the experimenter until the task was completed. The interactions were recorded with two head-mounted AKG C520 microphones to a 48 kHz stream with two separate channels and later on down sampled to 16 kHz for further signal processing. The recordings had an average length of ~15 min.

Subjects

Twenty-two speakers were recruited from the language talent study described above: twenty native speakers of German (10 female), rated either very high (10 talented speakers, 5 female) or very low on their phonetic talent (10 considerably less talented speakers, 5 female), age ranging from 20 to 42, and two further native speakers of English. The German native speakers all came from the greater region of Stuttgart in southern Germany, shared the same history of acquiring English (all started in fifth grade and no subject had stayed in an English-speaking country longer than 3 weeks), and were highly proficient in English. The grouping according to talent classes allowed us to test the influence of phonetic talent on the degree of phonetic convergence in spontaneous dialogs. The subjects were not informed about the research questions but simply asked to participate in a linguistic study involving reading and solving two spot-the-difference games with English dialog partners.

As per German law, psycholinguistic studies do not require consent by an ethics commission. All subjects were adults and were informed in detail in writing about their rights following the recommendations of ZENDAS (*Zentrale Datenschutzstelle der baden-württembergischen Universitäten*). All subjects then gave written and informed consent about their participation in the study in accordance with the Declaration of Helsinki, and have

received financial remuneration. No subjects took advantage of their right to have their data deleted.

Additionally two native speakers of English (male 33y, and female 57y) were recruited to participate in the dialogs. The male English native speaker spoke a General American (GA) accent—abbreviated as *AM*—and the female speaker a Standard Southern British English (SSBE) accent—abbreviated as *BR* in the results section. Male and female speakers were included in both native speaker groups to shed light on the as of yet inconclusive results on gender effects in phonetic convergence.

Amplitude Envelope Measurements

All target words (content words only) uttered by both speakers were manually extracted and labeled with the time markers—early (first third of the respective dialog) or late (last third)—to allow for tracking the speakers' potential increase in pronunciation similarity from an early to a late stage in the conversation. These cut-off points at the one-third and two-third markers thus, were set individually for each dialog, depending on its length. One reason for this was the assumption that a pair taking longer to complete the task and consequently also to finish the dialog, might also have taken an increased time adapting to one another (e.g., considering the establishment of common ground, usage of same or similar lexical items). In such a case, the first third of the dialog (= the “early” stage) would span a longer stretch of time, allowing for phonetic convergence to arise. The second reason was of a practical nature, namely the sufficient availability of content words uttered by both speakers within the early and late stages. This too could be aided by setting dynamic cut-off points dependent on individual dialog length.

The acoustic analysis of the data was based on slowly varying amplitude envelopes, comprising a smoothed global picture of the energy present in the signal (Wade et al., 2010; Lewandowski, 2012). The amplitude envelope analysis relies on the assumption that the information stored and used by humans in speech perception and production might be in parts represented in the form of envelopes—stored as linear time sequences. The previously manually extracted target words were thus stored as separate wav files, not normalized for length. The envelopes were then calculated by applying a standard Hilbert transform (built-in Matlab function). Such envelopes have been proven to contain information present in the auditory system, which also suffices to build intelligible speech as long as enough (usually two to three) frequency bands are available to provide an at least nominal spectral resolution (Shannon et al., 1995; Loizou et al., 1999; Wade et al., 2010). Here, four logarithmically-spaced frequency bands have been used. As the next step, the envelopes of the two words to be compared (tokens always stemming from the same type, e.g., “bread” speaker 1 vs. “bread” speaker 2) are passed to a cross-correlation function, which established their level of similarity (taking all four frequency bands separately). Returned was a single number: a *match value*—ranging between 0 and 1. The more similar the two amplitude envelopes, and consequently the pronunciations of the two words, the closer to 1 the match value. The lower the match value (approaching zero), the lower the similarity of the two envelopes (full script for reimplementation purposes available in Lewandowski, 2012).

Amplitude envelopes can be viewed as “representations that more faithfully encode the speech signal as it unfolds over time without making specific assumptions about what types of cues might be extracted or which regions of the signal are the most important” (Wade et al., 2010, p. 231). As it might be a very individual matter where precisely in the signal convergence takes place in the utterances of a speaker, amplitude envelopes are able to capture changes across the whole spectrum as it unfolds in time and not only for one specific feature, as VOT or (a number of) vowel formants. The method does not call for any front-end analyses and comes with a great transparency and compact output form, which are further advantages for convergence measurements (Wade et al., 2010). Furthermore, it captures spectral information present all over the signal, and since no timing normalization is performed neither on the raw word signals nor the extracted envelopes, the subsequent comparison of envelopes also accounts for timing information in the signals (e.g., lengthened or shortened segments receive lower match values). Compared to previous studies on imitation and convergence, such a method avoids single-feature tracking (such as individual formants or VOT measurements), and at the same time allows an analysis at a relatively fine-grained (word) level, instead of resorting to longer chunks which might deliver less reliable similarity estimations.

ANALYSIS AND RESULTS

The analysis of the convergence effects of the German speakers in the current study does not follow Lewandowski (2012) but has been completely re-calculated employing linear mixed effects modeling. This allowed the inclusion of by-speaker and by-partner random intercepts and slopes, which are extremely useful for handling datasets heavily loaded with individual differences and, at the same time, rather small effect sizes to be expected, which are most often found in research into phonetic adaptation. All statistical analyses were conducted using R version 3.4.4. Data manipulation was performed with the *tidyverse* package (Wickham, 2017), the linear mixed effects modeling was conducted with *lmerTest* (Kuznetsova et al., 2017) and the *leaps* package (Lumley and Miller, 2017). Data visualizations were created using *ggplot2* (Wickham, 2009) package. The *lmerTest* package automatically includes the significance values for the coefficients based on t-tests using the Satterthwaite’s method. Since the variables stemming from the psychological and cognitive tests were based on very different scales, the variables were z-scored by test prior to analysis to ensure comparability of effects across the different scaling systems used for the psychological and cognitive measures.

Phonetic Convergence and Language Talent

The analysis presented in this section involves match values calculated for the following *dialog times*: *early*—items (words) from an early point in the dialog from both speaking partners and *late*—again, early items from the English native speaker but paired with late items from the German subject. *The early*

set (Time 1) thus defines how similar the two interacting partners’ pronunciations were in the beginning of the dialog, while *the late set* (Time 2) allows determining how similar the late pronunciation variants of the German participants were compared to the previously heard English items. An expected finding would be an effect for Time – indicating that the similarity of the dialog partners’ envelopes changed during the conversation (taken the experimental group altogether). Should Time and Talent interact, this would point to differences in the envelope similarity changes between the talented and less talented speakers. Should *Gender* be included in the final model, it would speak for potential differences in convergence between male and female speakers in the study. The distribution of the predicted values (i.e., the match values) can be classified as unimodal (Hartigan’s dip test, $D = 0.05625$, $p > 0.05$) and has also been verified with an additional visual inspection of the corresponding density plot.

First, several random structures for the mixed model have been tested, selecting backwards from a maximal model (i.e., a model containing all variables; Crawley, 2013). The inclusion of a random slope of Talent*Time by Subject was not possible due to model convergence limits (number of random effects too high for number of cases observed). The model with a random slope with *Subject* over NS (NS = indicating which native speaker they talked to) was identified as the best fit, compared to a model with a random intercept for *Subject* only (Chisq 29.892, $\text{Pr(>Chisq)} = 3.229\text{e-}07^{***}$). The random and fixed effects of the final lmer model are given in Table 2. Following this, a linear mixed model was fitted with the match value as the dependent variable, and at first with an interaction of Time and Talent as the fixed factors only. Adding Gender as a further fixed factor in the next step proved to increase model fit significantly (**match value ~ Time*Talent + Gender + (1+NS|Subject)**, AIC –363.7, BIC –342.2, logLik 190.8, deviance –381.7, df.resid 71). On the other hand, adding the Native Speaker (NS) condition to the fixed structure, did not improve model fit (Chisq 1.0243, $\text{Pr(>Chisq)} = 0.3115$).

The factor Time alone was not significant in the model, indicating that the group as a whole did not show significant adaptation toward their dialog partners. However, as expected, the interaction of Time and Talent proved significant, (Table 3) indicating that the talented group (coded with “1”) converged more to their English native speaking partners from an early (Time 1) to a late point (Time 2) in the dialog than the less talented speakers (coded with “0”). Figure 2 shows the changes in match values from *early* to *late* separately for these two groups. The statistical analysis and the visualization reveal that the changes in pronunciation toward the conversational partners are of a subtle nature, as often reported in the literature (e.g., Pardo et al., 2017, 2018). Gender improved the model fit only when added as a simple factor but *not* in interaction with Time (model parameters with *Gender* in interaction with *Time*: AIC–362.73, BIC–338.91, model comparison with ANOVA: Chisq 1.0708, $\text{Pr(>Chisq)} = 0.3008$).

The fact that Gender appears only as a significant simple factor but not in interaction with Time indicates that although male subjects had on average higher match values than the female

TABLE 3 | Random and fixed effects for the linear mixed effects model: convergence in the dialogs and subjects' phonetic talent (significance values for the coefficients are based on *t*-tests using the Satterthwaite's method, provided within the *lmerTest* package).

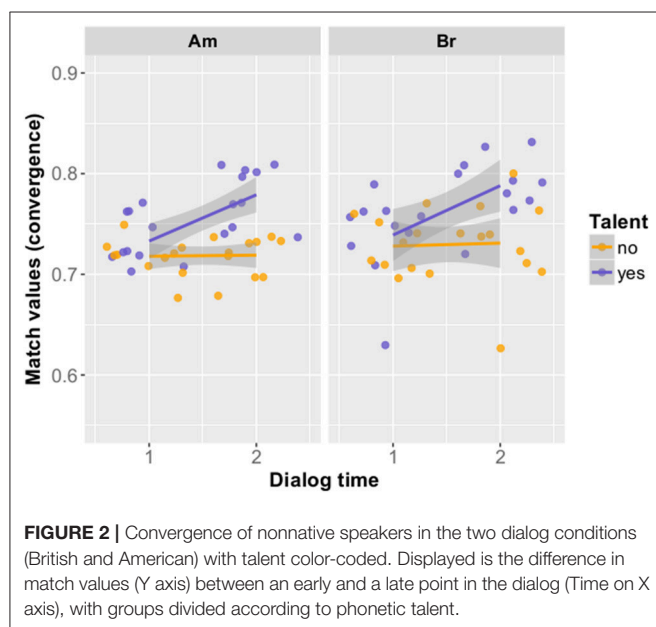
Random effects (observ.: 80, groups: Subject, 20)

Groups	Name	Variance	Std. Dev.	Corr
Subject	(Intercept)	0.0001259	0.01122	
	NS(British)	0.0014476	0.03805	−0.81
Residual		0.0002662	0.01632	

Fixed effects

	Estimate	Std. Error	df	<i>t</i> -value	Pr(> <i>t</i>)
(Intercept)	0.706259	0.008866	57.646855	79.658	<2e-16***
Time	0.002000	0.005159	40.000004	0.388	0.7003
Talent	−0.031730	0.012031	53.737784	−2.637	0.0109*
Gender	0.024640	0.004995	20.000000	4.932	8.03e-05***
Time:Talent	0.045500	0.007296	40.000004	6.236	2.20e-07***

p* < 0.05, **p* < 0.001.



participants throughout the experiment (i.e., irrespective of the dialog time), they did *not* converge significantly more than female talkers. Thus, the degree of convergence displayed within the dialogs was not influenced by gender—both male and female speakers behaved similarly. As to the general effect for male speakers displaying higher match values than female ones (i.e., not tied to convergence over time in the dialog), we cannot totally rule out the possibility of it simply being an artifact which is, underlyingly, *not* tied to gender at all (i.e., possibly arisen due to another feature that the male subjects might have shared amongst each other *or* with the English native speaking partners).

Summarizing the main findings of the convergence study, we can confirm that talented L2 learners converged phonetically

to their English native speaking partners in the course of the dialog; and, more interestingly, phonetically more gifted speakers did so to a significantly larger extent than the less talented subjects, proving that linguistic skill is a crucial factor for phonetic convergence to happen in an L2 environment. Gender, on the other hand, proved no significant factor for the degree of phonetic convergence in the current data set.

Phonetic Convergence and Psychological and Cognitive Factors

The second analysis involved the investigation of potential psychological and cognitive factors influencing the degree of phonetic convergence. Here, the group was treated irrespective of their phonetic talent since the constructs might operate partially interdependently (i.e., some cognitive subskills might feed into/surface as phonetic talent) and we are interested in obtaining a full picture of potential psychological and cognitive variables influencing each individual's phonetic adaptation (see also Lewandowski, 2013; Lewandowski et al., 2014).

Model Selection

In order to reduce the vast number of predictors, the subsequent analysis of the psychological variables was limited in the following cases. Instead of all subscales of the BAS test, only the mean BAS value was fed into the model (together with the BIS score). Instead of using all three scores from the E-Skala, only the two subscores (E-Skala Social and E-Skala Empathy) were considered since their sum forms the E-Skala General score, rendering it redundant. In contrast to this, all five factors from the NEO-FFI were considered, as they all represent very different personality aspects. For the Simon Test, the standard variable reflecting the *switch-cost* was included, being the difference between the reaction times for incongruent vs. congruent stimuli (SIMRTdiff). The measures for phonological working memory included in the models were digit span forward, digit span backward and non-word repetition span.

The predicted variable in the model was the amount of convergence between an early and a late point in the dialogs for every individual nonnative speaker (the difference between the match values at Time 2—*late*—and Time 1—*early*). Four subjects had to be excluded from the following analysis due to a lack of data points in one or more of the analyzed tests (Number of obs: 32, groups: Subject, 16). Model selection with the still considerably high number of model parameters was then aided with the *regsubsets* function from R's *leaps* package, a function assisting in regression subset selection in large datasets. Due to the high number of potential predictors and an exponentially higher number of possible combinations, a purely manual selection procedure was not possible at this stage. Since no other study to our knowledge included multiple personality and cognitive factors in a model of convergence, no clear hypotheses as to which factors will prove important for the prediction could be formulated a priori. Under these circumstances, neither a purely hypothesis-driven manual selection nor a fully manual testing procedure including all possible factor combinations was feasible. The chosen automatic selection aid was combined with a subsequent informed manual supervision and testing stage of

the suggested variable subset. The original model specification fed into the *regsubsets* function included: Convergence ~ open + agreeable + extroverted + neurotic + conscientious + BAS + BIS + EskalaEmp + EskalaSoc + WMnonword + WMfor + WMback + SIMRTdiff. *Regsubsets* also allows us to manually specify the selection algorithm used. In our case the *exhaustive algorithm* option, also called *brute force* algorithm, was chosen. It tests all possible combinations of variables and outputs the combination with the best fit for the dependent variable (maximizing Chi-square). The *regsubsets* function furthermore used the *nbest* setting of 1 (selecting one best model for each number of predictors) and the option *nvmax* of NULL (no limit on the number of variables that could form the best combination). Given these manual parameters the selection algorithm performed an *exhaustive* search and suggested a subset of 5 variables. These suggested variables were then manually fitted into the linear mixed model and their separate contributions to model fit were manually verified by performing anova model comparisons. As a consequence, one variable from the automatic preselection – *agreeableness* – was later on excluded after running an anova model comparison which proved it was not contributing to model fit (model incl. *agreeableness*: AIC 140.5, BIC=128.78, model ANOVA: Chisq 0.0015, Pr(>Chisq) 0.9694). The disparity between the subset selection and the final best model fit could arise due to missing data points for individual subjects' test scores within the data set. Not all subjects performed all psychological and cognitive tests, changing the number of available data points depending on the given model input.

The remaining four fixed factors for the model were *neurotic*, *open*, *BIS* and the *switch cost in the Simon Test* (SIMRTdiff). A linear mixed effects model specifying these fixed factors was then manually fitted using the *lmerTest* package. Collinearity was accounted for within the fixed factor correlation matrix and by additionally checking the variance inflation factors (vif). As no variable exceeded a vif score of 4, all four fixed factors were kept within the final model (AIC =142.5, BIC =132.2, logLik 78.3, deviance =156.5, df.resid 25; see **Table 4** for the random and fixed effects model parameters). The best random structure was achieved with *Subject* as a random intercept (Subject: Variance (Intercept) 2.537e-06, Std.Dev. 0.001593, Residual: Variance 4.375e-04, Std. Dev. 0.020917). Adding *Gender* in interaction with the remaining variables did not improve model fit (AIC=131.29, BIC=110.77, model ANOVA Chisq 2.7876, df 6, Pr(>Chisq) 0.835). Visual inspection of residual plots did not reveal any obvious deviations from normality. The native speaker condition (American or British) did not have a significant influence on the model fit. Neuroticism and openness had a positive impact on the degree of convergence in the dialogs, whereas the Behavior Inhibition score (BIS) and the switch costs in the Simon Test had a negative impact on convergence (see **Figures 3, 4**). In case of the latter this means that the lower the switch costs (RTincongruent - RTcongruent), meaning the better/faster someone was able to inhibit the wrong reaction, the more convergence there could be observed in the dialog task. For the BIS score it suggests that the less behaviorally inhibited the subject, the more did she or he converge in the dialog. Openness has the strongest impact

TABLE 4 | Random and fixed effects in the linear mixed model for convergence and the psychological and cognitive variables (significance values for the coefficients are based on *t*-tests using the Satterthwaite's method, provided within the *lmerTest* package).

Random effects

Groups	Name	Variance	Std. Dev.
Subject	(Intercept)	2.537e-06	0.001593
Residual		4.375e-04	0.020917

Fixed effects

	Estimate	Std. Error	df	t-value	Pr (> t)
(Intercept)	0.026120	0.003844	15.999998	6.795	4.70e-06***
Neurotic	0.016390	0.007068	15.999998	2.319	0.0340*
Open	0.027624	0.004047	15.999998	6.825	4.08e-06***
BIS	-0.018805	0.007309	15.999998	-2.573	0.0204*
Simon RTdiff	-0.010983	0.004102	15.999998	-2.677	0.0165*

p* < 0.05, **p* < 0.001.

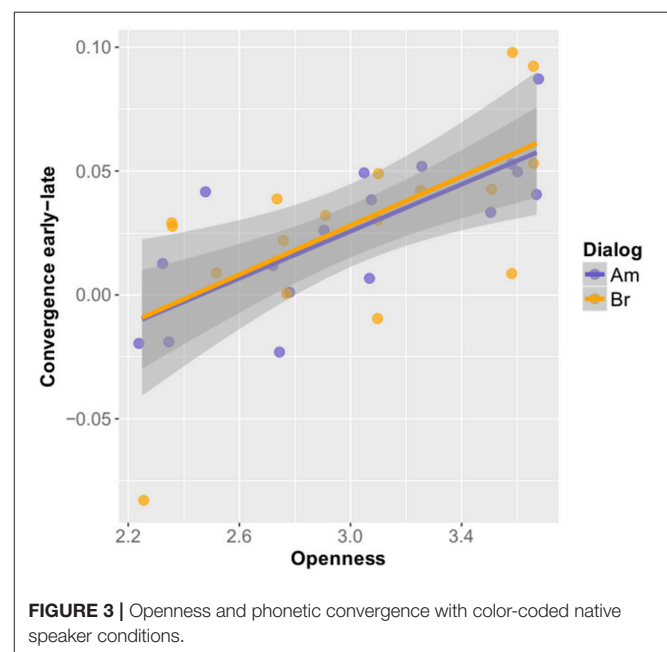
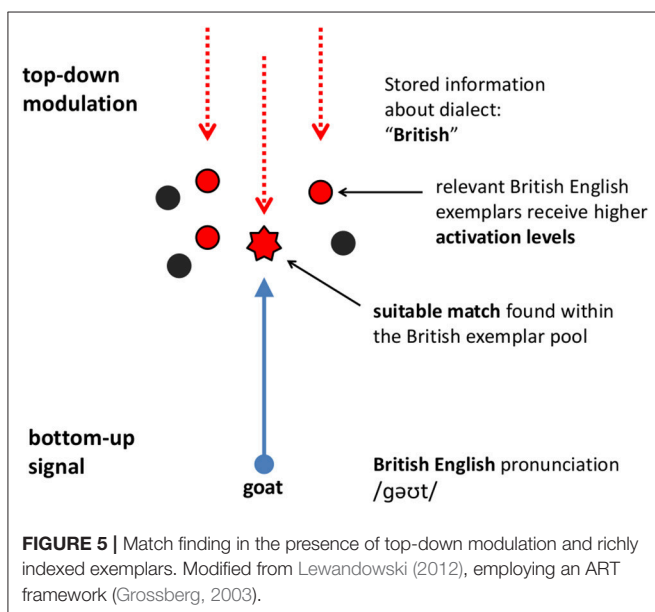
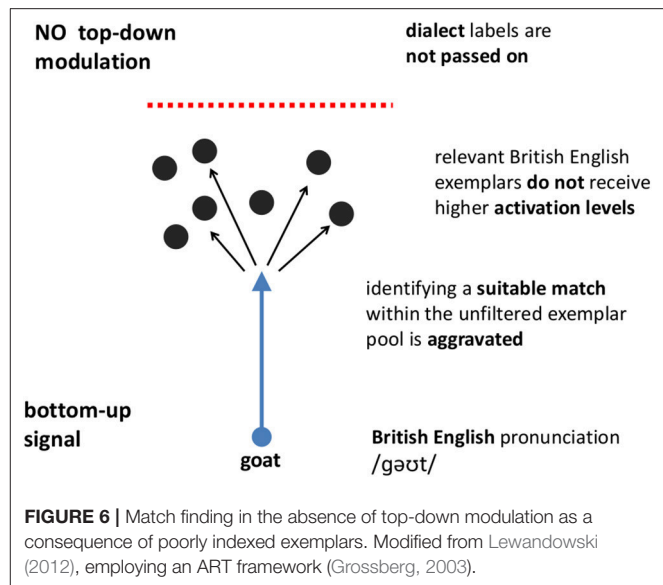
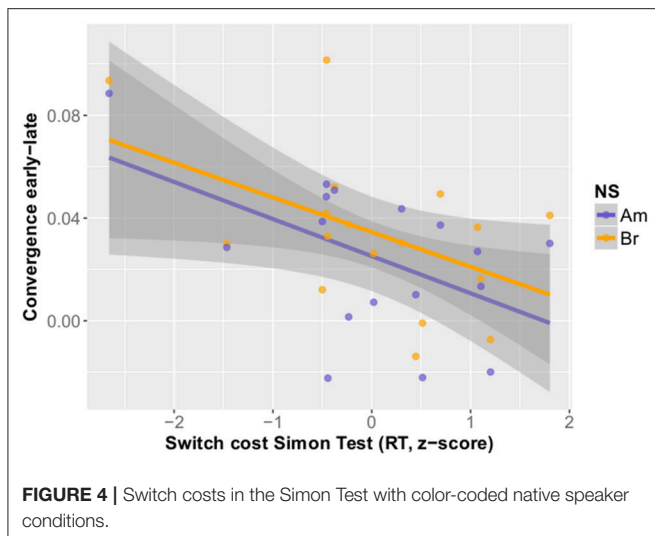


FIGURE 3 | Openness and phonetic convergence with color-coded native speaker conditions.

of all fixed factors on the convergence measure in the fitted mixed model.

DISCUSSION

A spontaneous dialog is influenced by a myriad of factors, all interacting and rapidly changing in time, requiring on-the-spot reactions from both speaking partners. Following such a conversation as an active listener-talker poses great attentional and memory, as well as executive demands. This is at the same time the *most natural* and *most complex* situation for a speaker to be in, and still, it is a situation with substantial individual differences arising. Repetition or shadowing designs undeniably profit from a far better control over the data by the experimenter.



It is certainly easier to make comparisons between specific phonetic features when the data is neat and simple, and not messy and complex as in natural conversations. However, the simpler and less complex the linguistic material, the less difficulty it poses for the subject and the better it can be controlled by her—as in perceiving isolated words or syllables. On the other hand, an interactive design featuring a conversation means more complexity to the signal and the necessity of arriving at sentence (or phrase) comprehension (including, a.o., lexical segmentation, an analysis of prosody and syntactic structure), which poses higher demands for long-term integration of information but also higher demands on attentional resources (Mattys, 1997; e.g., Davis and Johnsruide, 2007).

A listener-speaker has learned to draw from more than just the linguistic input to construct meaning and interpret it in a given situation and is used to having an overflow

of information (Davis and Johnsruide, 2007; Hawkins, 2010). These same listener-speakers are not only perfectly equipped for perception, they are also extremely flexible and capable of applying different styles that fit the current context, with an emphasis on the role of pronunciation. Phonology is said to be especially “fluid and skillfully deployed by individual speakers” to this end (Pavlenko and Blackledge, 2004, p. 244). Hawkins’ *Polysp system* assumes more specifically that *experienced* listeners use contextual information to identify the patterns of the current signal exactly according to the style and accent currently being used rather than just drawing from a “canonical” pool for pronunciation (Hawkins, 2003), which could lead to more accurate adaptation. We could show that it is not only experience but rather a certain talent for pronunciation that allows speakers to apply this flexibility for phonetic convergence in a dialog interaction.

Perceptual processes, forming the first step of adaptation, have been proposed to be interactive, incorporating top-down information, context and also expectations, instead of only the bottom-up signal (Grossberg, 2003; Davis and Johnsruide, 2007). Grossberg’s *Adaptive Resonance Theory* (2003), for instance, anticipates that the end-percept a listener obtains is emerging through the dynamic interplay of top-down memory-based information together with context-specific expectations and bottom-up (signal-driven) information (Figures 5, 6). The talented speakers in our study were significantly better at maneuvering between these signal types, and converging to their native speaking partners, proving to be more flexible speakers, who possibly pay more attention to the phonetic detail present in the speech of their partners, while maintaining an active conversation at the same time. Our current findings also suggest that the comprehensive approach for measuring pronunciation talent described by Jilka (2009a,b) is an accurate method for identifying phonetically skilled speakers, and to tear this skill apart from mere performance.

Language Talent and Personality

The claim that listeners are not able to notice and retain more phonetic detail in running speech than necessary for meaning construction, as put forward by Goldinger (2013), does not seem to hold for all language users. As in the case of our test subjects, L2 speakers identified as phonetically talented were well able to grasp fine phonetic detail within a conversation and re-use it shortly after to adapt to their conversation partners. The ability to converge phonetically, and effectively divide attentional resources between several tasks and skillfully master memory access, thus seems to be subject to individual differences. While talent for pronunciation certainly is not the only influencing factor in the complicated convergence mechanism, it is decisive for the eventual success of convergence since it seems to be operating within the core mechanism for spontaneous adaptation. This study contributed first insights into the still underinvestigated role psychological and cognitive individual differences play for phonetic convergence in conversational interaction.

First intuitions on what personality features characterize a good language learner, and potentially also a good “phonetic converger” went into the direction of extraversion, risk taking, lack of inhibition and self-esteem (MacIntyre and Noels, 1994; Hu and Reiterer, 2009). As suggested by Hu and Reiterer (2009), speakers characterized by higher openness might also be more open to the foreign social and cultural behaviors or ways of being, which could in turn lead to increased motivation and more success in acquiring a second language. In fact, the interpretation of the “openness to experience” dimension from the NEO-FFI used here is quite broad, encompassing, e.g., intelligence, imagination, perceptiveness, creativity, differentiated emotions, aesthetic sensitivity, need for variety, and unconventional values (Hu and Reiterer, 2009). Verhoeven and Vermeer (2002) also found moderate to high correlations between the personality trait openness and three subcomponents of communicative competence (organizational, pragmatic, and strategic competence) in young second language learners. The data from our current study suggest, in line with Yu et al.’s (2013) findings in their exposure experiment, that openness to experience is also an important factor in phonetic adaptation, possibly because it comes hand in hand with a greater perceptiveness and a higher communicative competence in general. One might speculate that greater imagination, for instance, might also positively influence not only having a more accurate theory of mind of our partner in a specific conversational situation but possibly also a greater understanding for speaker differences, including style and pronunciation. The lack of effects for extraversion in combination with convergence might in turn arise from the fact that openness taps more into communicative and partner-oriented skills than extraversion does. In Hu and Reiterer’s study there were no correlations between phonetic talent and openness nor extraversion. They argued that the nature of the phonetic tests (laboratory environment, closed task-sets, little free speech and no interaction with a partner) is in part responsible for these null findings since neither openness nor extraversion are strictly necessary skills to perform well in this kind of setting.

This explanation is in line with our finding, since openness to experience proved to be a factor for phonetic adaptation once the setting changed to an interactive one, resembling natural conversation.

As the core concept of the Behavior Inhibition Scale or — punishment sensitivity scale — is anxiety (Gray, 1981, 1982), and the interpretation includes all items referencing reactions to the anticipation of punishment (Hu and Reiterer, 2009), we can claim that a reduced anxiety or fear in anticipation of punishment has a positive impact on phonetic convergence. Furthermore, according to Gray and Braver (2002) individual differences on the BIS/BAS scale reflect IDs in approach and withdrawal – the two systems for action control, which are supposedly intertwined with cognitive aspects of control. This also sets them apart from other scales, such as the neuroticism/extraversion dimension (Carver et al., 2000). Translating this in terms of the current study, the less the L2 learners were in general prone to withdrawal in negative situations, the more positively they might have approached their conversational partner and embraced the situation of conversing in their L2, and also tried to (phonetically) adapt⁴ as well as possible in the L2 setting with no or little consideration of the consequences in case of a failure to do so. Given favorable constellations of other factors boosting convergence, they might have had a higher chance of succeeding in adaptation.

Neuroticism is interpreted as the individual difference in the way we experience distress and its subsequent impact on our behavioral and cognitive style (Hu and Reiterer, 2009). In our current study neuroticism had a positive impact on the degree of convergence. High neuroticism scores are usually tied to poorer performance in SLA (Gardner, 1985; Dörnyei, 2005) but can also have positive impacts, for example when observed in combination with a second dimension, such as extraversion. Robinson et al. (1994), for instance, report that students with high neuroticism and high extraversion scores performed better specifically on oral tests compared to written tests. It might thus make sense to interpret the result on neuroticism in conjunction with other traits favoring convergence, such as openness in our case. While it seems at first hard to establish a direct link between high neuroticism increasing the tendency to converge, it might still be possible if we relate it to dimensions investigated already in the first strand of research on Communication Accommodation Theory – as the need for social approval. Decreasing social distance is said to be one of the main socio-contextual factors for convergence (e.g., Giles and Powesland, 1975; Coupland, 2001; Giles, 2001). If an individual high on the neuroticism scale is heavily impacted by distress but at the same time has a high need for social approval, and (implicitly) knows that a lack of convergence could be interpreted by the dialog partner as diverging on purpose and have a negative impact on his or her evaluation, he or she might be inclined to strive for more adaptation in order to avoid being negatively evaluated and not belonging to the current in-group. To further evaluate

⁴Obviously, convergence cannot be controlled and steered to one’s liking, some aspects of it, however, remain under our control. See Lewandowski (2012) for a model proposal including aspects of control and consciousness.

the relation of neuroticism and convergence, additional tests involving the speakers' self-monitoring behavior and their need for social approval might be needed (e.g., Schweitzer et al., 2015).

Cognitive Skills

In contrast to the somewhat counterintuitive non-contribution of the working memory measures in the model of phonetic convergence, our inhibition measure (Simon Test) proved to be related to the degree of phonetic convergence, as expected. In contrast to Yu et al. (2013), who assessed their participants' attention switching skills based on an Autism Spectrum Quotient self-questionnaire (AQ, Baron-Cohen et al., 2001), attention inhibition in the current study was measured in an objective way, using a Simon Test. The smaller switch costs in the Simon Test (less time loss when switching between the congruent to the incongruent trials, Craft and Simon, 1970; Rota and Reiterer, 2009) proved to be positively related to the degree of phonetic convergence. The lower switch costs reflect a higher inhibitory control in the subjects, or, a higher mental flexibility. The more mentally flexible a speaker, the more likely he or she thus was to converge within the L2 dialog setting. The Simon Test is, in contrast to the Stroop Test, a purely nonverbal test which is based on spatial aspects/interference (with colorful dots appearing either on the left or the right side of the screen). In congruent trials the hand/finger with which the subject needs to click/press a button appears on the same side as the stimulus, in incongruent trials on the opposite side of the screen. The interference caused by the incongruent spatial information must be resolved by applying inhibition to the unwanted reaction, which leads to prolonged reaction times (= the switch costs). The negative result for working memory with respect to phonetic convergence, although rather unexpected, is in line with Hu et al.'s (2013) behavioral and neuroimaging findings which identified phonetic coding ability but not phonological working memory (PWM) to be predictive of L2 pronunciation aptitude in advanced learners. They argue that different learning stages require a different combination of skills – with PWM not being a distinctive factor between high and low aptitude advanced learners of a language. This could mean that PWM also contributes comparatively little to the phonetic convergence mechanism in the case of advanced L2 learners in contrast to, for instance, their inhibition skills. On the other hand, viewing inhibition in the framework of an *executive function* (EF) model (next to further EFs: attention switching and WM updating, e.g., Miyake et al., 2000), one also has to consider that the functions, albeit diverse, are never entirely separable. Inhibition and WM (updating) are intertwined for many tasks, including our chosen measure – the Simon Task. We cannot totally rule out the possibility that a portion of (P)WM's contribution to the degree of convergence was also—somewhat redundantly—captured within our inhibition measure. However, since none of the directly WM-related measures proved to be predictive, we can assume a relatively small contribution of WM compared to attention (inhibition).

Lev-Ari and Peperkamp (2014) used a manual version of the Stroop Test and the Simon Test (Craft and Simon, 1970) and correlated them to a lexical decision test for VOT-manipulated words in French. They report a stronger priming effect for

the VOT-manipulated words in perception for subjects with lower inhibition scores. Darcy et al. (2016) found significant correlations between a higher inhibition score in a retrieval-induced inhibition test (Lev-Ari and Peperkamp, 2013) and more accurate vowel perception as well as consonant production in their L2 learners, going into the direction of L1 performance. However, no relationship was found between vowel *production* and the inhibition score. Darcy et al. (2016) suggest that a more motor-based measure of inhibition, such as the Simon Test might have been better suited to capture production differences. As convergence inherently involves perception as well as production, this matches our finding of higher convergence paired with better inhibition skills measured with a Simon Test. Furthermore, our acoustic measure of amplitude envelope comparison is not limited to vowels nor consonant features but encompasses all acoustic features present in the signal at word level. Therefore, we might be able to more faithfully capture the influence of better inhibitory skills on L2 speech processing in our design, since it involves perception and production and a holistic acoustic measure.

A further interesting aspect emerges from Scharenborg et al.'s (2015) study on a perceptual learning effect and attention switch control. Subjects with worse attention-switching skills showed a stronger perceptual learning effect and were more prone to rely on lexical information in a word (= its orthography) than on the ambiguous sound of the target word. Listeners with higher attention-switching control attended more to the sound information and were not as easily inclined to adapt their phoneme categories to match the ambiguous sound. Listeners with a more efficient attention control are generally said to be better able to switch between bottom-up and top-down signals or weigh them more situation-appropriately (Mattys and Wiget, 2011; Mattys et al., 2013; Scharenborg et al., 2015). This could imply that subjects with a high inhibitory score (low switch costs) in our study might have paid more attention to the acoustic signal (bottom-up) and skillfully switched between lexical and acoustic properties, leaving enough room for handling the conversation and still picking up important and accurate acoustic-phonetic features of the dialog partner's speech to be re-used for their own productions – leading to significantly more phonetic convergence.

CONCLUSION

The current study provides evidence for the involvement of a specific phonetic talent component in successful phonetic convergence in an L2 dialog between German L2 learners of English and English native speakers, suggesting that in order for convergence to happen, the necessary phonetic skills are a prerequisite. Furthermore, several psychological and cognitive dimensions have been identified, which impact the degree to which nonnative speakers adapted their pronunciation in the dialogs. Amongst them, openness to experience seems to play an especially important role. This implies that further investigations into the reasons for phonetic convergence, especially in L2 contexts, should incorporate measures of phonetic skill as well

as personality and cognitive abilities into their designs, to control for any effects stemming from these individual differences within the speaker rather than from pure socially and/or contextually arising factors.

ETHICS STATEMENT

The study was approved by the local ethics committee from the University of Tübingen, since parts of it involved neurolinguistics measurements (a comparable permission is not needed for psycholinguistic studies as per German law). This study was carried out in accordance with the recommendations of ZENDAS (Zentrale Datenschutzstelle der baden-württembergischen Universitäten). All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved. All subjects were adults and have received financial remuneration. No subjects took advantage of their right to have their data deleted.

AUTHOR CONTRIBUTIONS

NL conducted the convergence study and the study linking phonetic convergence to the personality and cognitive measures including all analyses (sections Motivation and Goals, Main Study: Testing Phonetic Convergence in a Conversational Setting, Materials, Procedure, Subjects, Amplitude Envelope Measurements, Analysis and Results, Phonetic Convergence and Language Talent, Phonetic Convergence, and Psychological and Cognitive Factors, Discussion, and Conclusion). MJ granted access to the personality and psychological data set from the

Language Talent study and described the talent test battery for the purpose of this paper (sections Phonetic Convergence, Factors Influencing Phonetic Convergence, The Role of Phonetic Talent, Language Talent Study: Identification and Classification of Pronunciation Talent in Language Learners (Dogil and Reiterer, 2009), The Test Set-Up, Psychological and Cognitive Aspects, and Testing Phonetic Ability).

FUNDING

The study was supported by the German Research Council (DFG) through the DFG project Language talent and brain activity—the neural basis of pronunciation talent, the Collaborative Research Center SFB 732, project A4 Phonetic convergence in spontaneous speech (2010–2018), and the DFG project Prosodische Phrasierung in der auditorischen und visuellen Satzverarbeitung (project no. 159998143).

ACKNOWLEDGMENTS

This study is partially based on the data collected for the doctoral dissertation of Lewandowski (2012). An overview of the language talent study including the talent data has been already presented at the *New Sounds 2010* conference in Poznan (Jilka et al., 2010). The authors would like to thank Prof. Dr. Grzegorz Dogil (1952–2017) who was a driving force in both the Language Talent and the Brain project and as the supervisor of Lewandowski's dissertation (2012). He provided us with constant inspiration and support in our work and is deeply missed.

REFERENCES

- Abrego-Collier, C., Grove, J., Sonderegger, M., and Yu, A. C. L. (2011). "Effects of speaker evaluation on phonetic convergence," in *Proceedings of ICPhS XVII* (Hong Kong), 192–195.
- Abu-Rabia, S., and Kehat, S. (2004). The critical period for second language pronunciation: is there such a thing? *Edu. Psychol.* 24, 77–97. doi: 10.1080/0144341032000146467
- Aguilar, L., Downey, G., Krauss, R., Pardo, J. S., Lane, S., and Bolger, N. (2016). A dyadic perspective on speech accommodation and social connection: both partners' rejection sensitivity matters: a dyadic perspective on speech. *J. Pers.* 84, 165–177. doi: 10.1111/jopy.12149
- Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, M. G., Garrod, S., et al. (2001). The HCRC map task corpus. *Lang. Speech* 34, 351–366. doi: 10.1177/002383099103400404
- Babel, M. (2012). Evidence for phonetic and social selectivity in phonetic accommodation. *J. Phon.* 40, 177–189. doi: 10.1016/j.wocn.2011.09.001
- Babel, M. E. (2009). *Phonetic and Social Selectivity in Speech Accommodation*. University of California, Berkeley.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201
- Bailly, G., and Lelong, A. (2010). "Speech dominoes and phonetic convergence", in *Proceedings of INTERSPEECH 2010*, 1153–1156.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/A:1005653411471
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bongaerts, T., Planken, B., and Schils, E. (1995). "Can late starters attain a native accent in foreign language? a test of the critical period hypothesis," in *The Age Factor in Second Language Acquisition*, eds D. Singleton and Z. Lengyel (Clevedon: Multilingual Matters), 30–50.
- Borkenau, P., and Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae: Handanweisung*. Hogrefe Verlag für Psychologie.
- Bradlow, A. R., Baker, R. E., Choi, A., Kim, M., and van Engen, K. J., (2007). The wildcat corpus of native- and foreign-accented english. *J. Acoust. Soc. Am.* 121:3072. doi: 10.1121/1.4808527
- Brouwer, S., Mitterer, H., and Huettig, F. (2010). Shadowing reduced speech and alignment. *J. Acoust. Soc. Am.* 128, EL32–EL37. doi: 10.1121/1.3448022
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Lang. Var. Change* 14, 261–290. doi: 10.1017/S0954394502143018
- Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language*, 82, 711–733.
- Cappella, J. N., and Planalp, S. (1981). Talk and silence sequences in informal conversations III: interspeaker influence. *Hum. Commun. Res.* 7, 117–132. doi: 10.1111/j.1468-2958.1981.tb00564.x
- Carroll, J. B. (1981). "Twenty-Five years of research on foreign language aptitude," in *Individual Differences and Universals in Language Learning Aptitude*, ed K. S. Diller (Rowley, MA: Newbury House), 83–118.
- Carroll, J. B., and Sapon, S. M. (1959). *Modern Language Aptitude Test (MLAT)*. San Antonio, TX: Psychological Corporation.
- Carver, C. S., Sutton, S. K., and Scheier, M. F. (2000). Action, emotion, and personality: emerging conceptual integration. *Personal. Soc. Psychol. Bull.* 26, 741–751. doi: 10.1177/0146167200268008
- Carver, C. S., and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J. Pers. Soc. Psychol.* 67, 319–333. doi: 10.1037/0022-3514.67.2.319

- Conboy, B. T., Sommerville, J. A., and Kuhl, P. K. (2008). Cognitive control factors in speech perception at 11 months. *Dev. Psychol.* 44, 1505–1512. doi: 10.1037/a0012975
- Costa, P. T., and McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and Neo Five Factor Inventory: Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Coupland, N. (2001). “Language, Situation, and the Relational Self: Theorizing Dialect-Style in Sociolinguistics,” in *Style and Sociolinguistic Variation*, eds P. Eckert and J. R. Rickford (Cambridge: Cambridge University Press), 185–210.
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51, 42–100. doi: 10.1016/j.cogpsych.2004.12.001
- Craft, J. L., and Simon, J. R. (1970). Processing symbolic information from a visual display: interference from an irrelevant directional cue. *J. Exp. Psychol.* 83(3 Pt.1), 415–420. doi: 10.1037/h0028843
- Crawley, M. J. (2013). *The R book, 2nd Edn*. Chichester: Wiley.
- Darcy, I., Mora, J. C., and Daidone, D. (2016). The Role of inhibitory control in second language phonological processing: inhibitory control and L2 phonology. *Lang. Learn.* 66, 741–773. doi: 10.1111/lang.12161
- Davies, A. (2003). *The Native Speaker: Myth and Reality*. Clevedon: Multilingual Matters. Available online at: <http://www.loc.gov/catdir/toc/fy035/2002014538.html>
- Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147. doi: 10.1016/j.heares.2007.01.014
- De Looze, C., Oertel, C., Rauzy, S., and Campbell, N. (2011). “Measuring dynamics of mimicry by means of prosodic cues in conversational speech,” in *International Conference on Phonetic Sciences (ICPhS)* (Hong Kong), 1294–1297.
- Delvaux, V., and Soquet, A. (2007). “Inducing imitative phonetic variation in the laboratory,” in *Proceedings of the 16th International Conference of Phonetic Sciences* (Saarbrücken), 369–372.
- Dias, J. W., and Rosenblum, L. D. (2011). Visual influences on interactive speech alignment. *Perception* 40, 1457–1466. doi: 10.1068/p7071
- Dogil, G., and Reiterer, S. M. (eds.). (2009). *Language Talent and Brain Activity, Vol. 1. Trends in Applied Linguistics*. Berlin: De Gruyter. doi: 10.1515/9783110215496
- Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Engen, K. J., van, Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., and Bradlow, A. R. (2010). The wildcat corpus of native-and foreign-accented english: communicative efficiency across conversational dyads with varying language alignment profiles. *Lang. Speech* 53, 510–540. doi: 10.1177/0023830910372495
- Flege, J., and Hillenbrand, J. (1987). “Limits on phonetic accuracy in foreign language speech production,” in *Interlanguage Phonology*, eds G. Ioup and S. Weinberger (Cambridge, MA: Newbury House), 176–203.
- Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*. London: Edward Arnold.
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., and Emslie, H. (1994). The Children’s test of nonword repetition: a test of phonological working memory. *Memory* 2, 103–127. doi: 10.1080/09658219408258940
- Giles, H. (2001). “Couplandia and beyond,” in *Style and Sociolinguistic Variation*, eds P. Eckert, and J. R. Rickford (Cambridge: Cambridge University Press), 211–19.
- Giles, H., and Ogay, T. (2006). “Communication accommodation theory,” in *Explaining Communication: Contemporary Theories and Exemplars*, eds B. B. Whaley, and W. Samter (Mahwah, NJ: Lawrence Erlbaum Associates), 293–310.
- Giles, H., and Powesland, P. F. (1975). *Speech Style and Social Evaluation, Vol. 7. European Monographs in Social Psychology*. London: Academic Press.
- Goldinger, S. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *J. Exp. Psychol.* 22, 1166–1183. doi: 10.1037//0278-7393.22.5.1166
- Goldinger, S. (1998). Echoes of echoes? an episodic theory of lexical access. *Psychol. Rev.* 105, 251–279. doi: 10.1037/0033-295X.105.2.251
- Goldinger, S. D. (2013). The cognitive basis of spontaneous imitation: evidence from the visual world. *Proc. Mtgs. Acoust.* 19:060136. doi: 10.1121/1.4800039
- Gollan, T. H., Sandoval, T., and Salmon, D., P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychol. Sci.* 22, 1155–1164. doi: 10.1177/0956797611417002
- Gray, J. A. (1981). “A critique of Eysenck’s theory of personality,” in *A Model for Personality*, ed H. J. Eysenck (Berlin: Springer-Verlag), 246–276.
- Gray, J. A. (1982). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. New York, NY: Oxford University Press.
- Gray, J. R., and Braver, T. S. (2002). Personality predicts working-memory-related activation in the caudal anterior cingulate cortex. *Cognit. Affect. Behav. Neurosci.* 2, 64–75. doi: 10.3758/CABN.2.1.64
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingual. Lang. Cognit.* 1:67. doi: 10.1017/S1366728998000133
- Gregory, S. W. (1983). A quantitative analysis of temporal symmetry in microsocial relations. *Am. Sociol. Rev.* 48, 129–135. doi: 10.2307/2095151
- Gregory, S. W. (1986). Social psychological implications of voice frequency correlations: analyzing conversation partner adaption by computer. *Soc. Psychol. Quart.* 49, 237–246. doi: 10.2307/2786806
- Gregory, S. W., and Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *J. Pers. Soc. Psychol.* 70, 1231–1240. doi: 10.1037/0022-3514.70.6.1231
- Grossberg, S. (2003). Resonant neural dynamics of speech perception. *J. Phon.* 31, 423–445. doi: 10.1016/S0095-4470(03)00051-2
- Guiora, A. (1990). A psychological theory of second language pronunciation. *Toegepaste Taalwetenschap in Artikelen* 37, 15–23. doi: 10.1075/ttwia.37.02gui
- Hari, R., and Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends Cogn. Sci.* 5, 525–532. doi: 10.1016/S1364-6613(00)01801-5
- Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *J. Phon.* 31, 373–405. doi: 10.1016/j.wocn.2003.09.006
- Hawkins, S. (2010). Phonological features, auditory objects, and illusions. *J. Phon.* 38, 60–89. doi: 10.1016/j.wocn.2009.02.001
- Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., and Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: behavioural predictors and neural substrates. *Brain Lang.* 127, 366–376. doi: 10.1016/j.bandl.2012.11.006
- Hu, X., and Reiterer, S. M. (2009). “Personality and pronunciation talent in second language acquisition,” in *Language Talent and Brain Activity Reiterer Trends in Applied Linguistics*, eds G. Dogil and S. M. (Berlin: de Gruyter), 97–129.
- Jilka, M. (2000). The Contribution of Intonation to the Perception of Foreign Accent. Doctoral Dissertation, Arbeiten des Instituts für Maschinelle Sprachverarbeitung (AIMS) Vol. 6, University of Stuttgart.
- Jilka, M. (2009a). “Assessment of phonetic ability,” in *Language Talent and Brain Activity Trends in Applied Linguistics*, eds G. Dogil and S. M. Reiterer (Berlin: de Gruyter), 17–66.
- Jilka, M. (2009b). “Talent and proficiency in language,” in *Language Talent and Brain Activity Trends in Applied Linguistics*, eds G. Dogil and M. S. Reiterer (Berlin: de Gruyter), 1–16.
- Jilka, M., Lewandowska, N., and Rota, G. (2010). “Investigating the concept of talent in phonetic performance,” in *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010*, eds D. K. Katarzyna, M. Wrembel, and M. Kul (Poznan).
- Johnson, K. (1997). “Speech perception without speaker normalization: an exemplar model,” in *Talker Variability in Speech Processing*, eds K. Johnson and J. W. Mullennix (San Diego, CA: Academic Press), 145–165.
- Juszyk, P. W. (2002). How infants adapt speech-processing capacities to native-language structure. *Curr. Dir. Psychol. Sci.* 11, 15–18. doi: 10.1111/1467-8721.00159
- Kelso, S. J. A. (1997). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: A Bradford book. MIT Press.
- Kelso, S. J. A., and Engström, D. A. (2006). *The Complementary Nature*. Cambridge, MA: MIT Press.
- Kim, M., Horton, W. S., and Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Lab. Phonol.* 2, 125–126. doi: 10.1515/labphon.2011.004
- Krauss, R. M., and Pardo, J. S. (2004). Is alignment always the result of automatic priming? *Behav. Brain Sci.* 27, 203–204. doi: 10.1017/S0140525X0436005X

- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest Package: tests in linear mixed effects models. *J. Statist. Software* 82, 1–26. doi: 10.18637/jss.v082.i13
- Lallier, M., M.-Tainturier, J., Dering, B., Donnadieu, S., Valdois, S., and Thierry, G. (2010). Behavioral and ERP evidence for amodal sluggish attentional shifting in developmental dyslexia. *Neuropsychologia* 48, 4125–4135. doi: 10.1016/j.neuropsychologia.2010.09.027
- Leibetseder, M., Laireiter, A.-R., Riepler, A., and Köller, T. (2001). E-Skala: fragebogen zur erfassung von empathie - beschreibung und psychometrische eigenschaften. *Zeitschr. Differ. Diagn. Psychol.* 22, 70–85. doi: 10.1024/0170-1789.22.1.70
- Lev-Ari, S., and Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *J. Phon.* 41, 320–331. doi: 10.1016/j.wocn.2013.06.002
- Lev-Ari, S., and Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception. *J. Phon.* 47, 36–46. doi: 10.1016/j.wocn.2014.09.001
- Levitan, R., Benus, S., Gravano, A., and Hirschberg, J. (2015). "Entrainment and turn-taking in human-human dialogue," in *AAAI 2015 Spring Symposium on Turn-Taking and Coordination in Human-Machine Interaction* (Palo Alto, CA: The AAAI Press).
- Lewandowski, N. (2012). *Talent in Nonnative Phonetic Convergence*. Doctoral Dissertation, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.
- Lewandowski, N. (2013). "Phonetic convergence and individual differences in non-native dialogs," in *Book of Abstracts New Sound Conference 2013, Montréal*. (Montréal, QC)
- Lewandowski, N., and Duran, D. (2018). "Testing speech perception today and tomorrow: serious computer games as perception tests," in *Proceedings of the 29. Conference of Electronic Signal Processing (ESSV) in Ulm. Studententexte zur Sprachkommunikation, no. 90*, eds A. Berton, U. Haiber, W. Minker (Ulm).
- Lewandowski, N., Schweitzer, A., Duran, D., and Dogil, G. (2014). "An exemplar-based hybrid model of phonetic adaptation," in *Paper presented at the Georgetown University Round Table in Linguistics (GURT 2014): Usage-based Approaches to Language, Language Learning, and Multilingualism in Washington D.C.*
- Linck, J. A., Schwieter, J. W., and Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingual. Lang. Cognit.* 15, 651–662. doi: 10.1017/S136672891100054X
- Loizou, P. C., Dorman, M., and Tu, Z. (1999). On the number of channels needed to understand speech. *J. Acoust. Soc. Am.* 106, 2097–2103. doi: 10.1121/1.427954
- Lumley, T., and Miller, A. (2017). leaps: Regression Subset Selection. R package version 3.0. Available online at: <https://CRAN.R-project.org/package=leaps> (accessed February 24, 2019).
- MacIntyre, P. D., and Noels, K. A. (1994). The good language learner: a retrospective review. *System* 22, 269–280. doi: 10.1016/0346-251X(94)90062-0
- Markham, D. (1997). *Phonetic Imitation, Accent, and the Learner*. Ph.D. Dissertation, Travaux de l'Institut de Linguistique de Lund 33. Lund University Press.
- Matarazzo, J. D., Weitman, M., Saslow, G., and Wiens, A. N. (1963). Interviewer influence on durations of interviewee speech. *Verb. Learn. Verb. Behav.* 1, 451–458.
- Mattys, S. L. (1997). The use of time during lexical processing and segmentation: a review. *Psychon. Bull. Rev.* 4, 310–329. doi: 10.3758/BF03210789
- Mattys, S. L., Seymour, F., Attwood, A. S., and Munafò, M. R. (2013). Effects of acute anxiety induction on speech perception: are anxious listeners distracted listeners? *Psychol. Sci.* 24, 1606–1608. doi: 10.1177/0956797612474323
- Mattys, S. L., and Wiget, L. (2011). Effects of cognitive load on speech recognition. *J. Mem. Lang.* 65, 145–160. doi: 10.1016/j.jml.2011.04.004
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Moyer, A. (1999). Ultimate attainment in L2 phonology: the critical factors of age, motivation, and instruction. *Stud. Second Lang. Acquisit.* 21, 81–108. doi: 10.1017/S0272263199001035
- Namy, L. L., Nygaard, C. L., and Sauerteig, D. (2002). Gender differences in vocal accommodation: the role of perception. *J. Lang. Soc. Psychol.* 21, 422–432.
- Natale, M. (1975a). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *J. Pers. Soc. Psychol.* 32, 790–804. doi: 10.1037//0022-3514.32.5.790
- Natale, M. (1975b). Social desirability as related to convergence of temporal speech patterns. *Percept. Mot. Skills* 40, 827–830. doi: 10.2466/pms.1975.40.3.827
- Neufeld, G. (1987). "On the acquisition of prosodic and articulatory features in adult language learning". In *Interlanguage Phonology*, eds G. Ioup, and S. Weinberger, (Cambridge, MA: Newbury House), 321–332.
- Nielsen, K. (2007). 'Implicit phonetic imitation is constrained by phonemic contrast,' in *Proceedings of the 16th International Conference of Phonetic Sciences*, eds J. Trouvain and W. J. Barry (Saarbrücken), 1961–1964.
- Nielsen, K. (2008). *Word-Level and Feature-Level Effects in Phonetic Imitation*. Los Angeles, CA: University of California.
- Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *J. Phon.* 39, 132–142. doi: 10.1016/j.wocn.2010.12.007
- Ou, J., and Law, S.-P. (2017). Cognitive basis of individual differences in speech perception, production and representations: the role of domain general attentional switching. *Attent. Percept. Psychophys.* 79, 945–963. doi: 10.3758/s13414-017-1283-z
- Ou, J., S.-Law, P., and Fung, R. (2015). Relationship between individual differences in speech processing and cognitive functions. *Psychon. Bull. Rev.* 22, 1725–1732. doi: 10.3758/s13423-015-0839-y
- Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *J. Psycholinguist. Res.* 5, 261–283. doi: 10.1007/BF01067377
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* 119:2382–2393. doi: 10.1121/1.2178720
- Pardo, J. S., Gibbons, R., Suppes, A., and Krauss, R. M. (2012). Phonetic convergence in college roommates. *J. Phon.* 40, 190–197. doi: 10.1016/j.wocn.2011.10.001
- Pardo, J. S., Jay, I. C., Hoshino, R., Hasbun, S. M., Sowemimo-Coker, C., and Krauss, R. M. (2013). Influence of role-switching on phonetic convergence in conversation. *Discourse Process.* 50, 276–300. doi: 10.1080/0163853X.2013.778168
- Pardo, J. S., Urmanche, A., Wilman, S., and Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attent. Percept. Psychophys.* 79, 637–659. doi: 10.3758/s13414-016-1226-0
- Pardo, J. S., Urmanche, A., Wilman, S., Wiener, J., Mason, N., Francis, K., et al. (2018). A comparison of phonetic convergence in conversational interaction and speech shadowing. *J. Phon.* 69, 1–11. doi: 10.1016/j.wocn.2018.04.001
- Park, J.-E. (2007). Co-construction of nonnative speaker identity in cross-cultural interaction. *Appl. Linguist.* 28, 339–360. doi: 10.1093/applin/amm001
- Pavlenko, A., and Blackledge, A. (2004). *Negotiation of Identities in Multilingual Contexts, Vol. 45. Bilingual Education and Bilingualism*. (Clevedon: Multilingual Matters). Available online at: <http://www.loc.gov/catdir/toc/ecip042/2003008658.html>
- Pickering, M. J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–190. doi: 10.1017/S0140525X04000056
- Pickering, M. J., and Garrod, S. (2005). "Automaticity of language production in monologue and dialogue," in *Automaticity and Control in Language Processing*, eds A. Meyer, L. Wheeldon, and A. Krott (Hove: Psychology Press).
- Pickering, M. J., and Garrod, S. (2006). Alignment as the basis for successful communication. *Res. Lang. Comput.* 4, 203–228. doi: 10.1007/s11168-006-9004-0
- Pickering, M. J., and Garrod, S. (2013). Forward models and their implications for production, comprehension, and dialogue. *Behav. Brain Sci.* 36, 377–392. doi: 10.1017/S0140525X12003238
- Pierrehumbert, J. B. (2006). The next toolkit. *J. Phon.* 34, 516–530.
- Pitts, M. J., and Giles, H. (2008). "Social psychology and personal relationships: accommodation and relational influence across time and contexts," in *Handbook of Interpersonal Communication, Vol. 2*, eds G. Antos, E. Ventola, T. Weber, and K. Knapp (Berlin: Mouton de Gruyter), 15–31.
- Robinson, D., Gabriel, N., and Katchan, O. (1994). Personality and second language learning. *Pers. Individ. Dif.* 16, 143–157. doi: 10.1016/0191-8869(94)90118-X

- Robinson, P. (2003). "Attention and memory during SLA," in *The Handbook of Second Language Acquisition*, eds C. J. S. Doughty and M. H. Long (Malden, MA: Blackwell), 631–678.
- Rota, G., and Reiterer, S. M. (2009). "Cognitive aspects of pronunciation talent. in *Language Talent and Brain Activity Trends in Applied Linguistics*, eds G. Dogil and S. M. Reiterer (Berlin: de Gruyter), 67–96.
- Safronova, E. (2016). *The Role of Cognitive Ability in the Acquisition of Second Language Perceptual Phonological Competence*. Dissertation. Barcelona: Universitat de Barcelona.
- Scharenborg, O., Weber, A., and Janse, E. (2015). The role of attentional abilities in lexically guided perceptual learning by older listeners. *Attent. Percept. Psychophys.* 77, 493–507. doi: 10.3758/s13414-014-0792-2
- Schneiderman, E., and Desmarais, C. (1988). "a neuropsychological substrate for talent in second language acquisition," in *The Exceptional Brain*, ed L. K. Obler (New York, NY: Guilford Press), 103–26.
- Schweitzer, A., and Lewandowski, N. (2013). "Convergence of articulation rate in spontaneous speech," in *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech, Lyon)*, 525–529.
- Schweitzer, A., and Lewandowski, N. (2014). "Social factors in convergence of F1 and F2 in spontaneous speech," in *Proceedings of the 10th International Seminar on Speech Production (Cologne)*.
- Schweitzer, A., Lewandowski, N., and Duran, D. (2015). "Attention, Please! Expanding the GECO Database," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, edited by The Scottish Consortium for ICPhS 2015. (Glasgow) Available online at: <http://www.icphs2015.info/pdfs/Papers/ICPHS0620.pdf>
- Serafini, E. J. (2017). Exploring the dynamic long-term interaction between cognitive and psychosocial resources in adult second language development at varying proficiency. *Modern Lang. J.* 101, 369–390. doi: 10.1111/modl.12400
- Shannon, R. V., F.-G., Zeng, Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303
- Simard, L., Taylor, D., and Giles, H. (1976). Attribution processes Ad interpersonal accommodation in a bilingual setting. *Lang. Speech* 19, 374–387. doi: 10.1177/002383097601900408
- Simon, J. R. (1990). "The effects of an irrelevant directional cue on human information processing" in *Stimulus-Response Compatibility: An Integrated Perspective*, eds R. W. Proctor and T. G. Reeve (Amsterdam: North Holland Publishing Company), 31–88.
- Skehan, P. (2003). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Smith, C. L. (2007). "Prosodic accommodation by french speakers to a non-native interlocutor," in *Proceedings of the 16th International Conference of Phonetic Sciences*. (Saarbrücken), 1081–84.
- Street, R. (1984). Speech convergence and speech evaluation in fact-finding interviews. *Hum. Commun. Res.* 11, 139–169. doi: 10.1111/j.1468-2958.1984.tb00043.x
- Strobel, A., Beauducel, A., Debener, S., and Brocke, B. (2001). Eine deutschsprachige version des BIS/BAS-fragebogens von carver und white. *Zeitschr. Differ. Diagn. Psychol.* 22, 216–227. doi: 10.1024//0170-1789.22.3.216
- Styles, E. A. (2006). *The Psychology of Attention, 2nd Edition*. Hove: Psychology Press.
- Vais, J., Lewandowski, N., and Walsh, M. (2015). *Investigating Frequency of Occurrence Effects in L2 Speakers: Talent Matters*. In *ICPhS XVIII*, paper no. 0723. Glasgow. Available online at: <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0723.pdf>
- Verhoeven, L., and Vermeer, A. (2002). Communicative competence and personality dimensions in first and second language learners. *Appl. Psycholinguist.* 23, 361–374. doi: 10.1017/S014271640200303X
- Wade, T., Dogil, G., Schütze, H., Walsh, M., and Möbius, B. (2010). Syllable frequency effects in a context-sensitive segment production model. *J. Phon.* 38, 227–239. doi: 10.1016/j.wocn.2009.10.004
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.
- Wickham, H. (2017). *tidyverse: Easily Install and Load the 'Tidyverse.'* R package version 1.2.1. Available online at: <https://CRAN.R-project.org/package=tidyverse> (accessed February 24, 2019).
- Willemyns, M., Gallois, C., Callan, V., and Pittam, J. (1997). Accent accommodation in the job interview: the impact of interviewer accent and gender. *J. Lang. Soc. Psychol.* 16, 3–22. doi: 10.1177/0261927X970161001
- Yu, A. C., Abrego-Collier, C., and Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: subjective attitude, personality and 'Autistic' traits. *PLoS ONE* 8:e74746. doi: 10.1371/journal.pone.0074746

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Lewandowski and Jilka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.