A Phonotactic Grammar for Perception

Tom Lentz, René Kager

1. Phonotactic knowledge and its effects on perception

The phonotactic status of sound combinations has been shown in various ways to influence speech perception, suggesting that knowledge of phonotactics is represented in the minds of listeners. Phonotactic knowledge has the potential to help correct errors and resolve ambiguities in the acoustic signal (e.g. Hallé et al. 1998) and can also provide other valuable information, for instance cues to detect word boundaries. As phonotactic influences on perception depend on the listener's language (e.g. Dupoux et al. 1999, Weber and Cutler 2006; Lentz 2011: Ch. 3-4), phonotactic knowledge also has to be language-specific and cannot be assumed to be reduced to phonetics proper.

This contribution discusses the architecture of grammar that can generate the information necessary to explain perceptual effects of phonotactic well-formedness. The grammar captures both probabilistic and categorical phonotactic knowledge and describes both its categorical and gradient effects on speech perception.

The simplest understanding of phonotactic knowledge is that it is knowledge of distributions of sound combinations. This understanding requires no assumptions about complex and/or phonotactic-specific learning predisposition. However, not all phonotactic knowledge can be directly related to mere distributional knowledge: some unattested combinations are treated as more well-formed than others (Berent et al. 2007, Moreton 2002). Listeners must be assumed to (also) have more complex phonotactic knowledge than simple tabulations of frequency counts. Nevertheless, many phonotactic effects on perception are closely correlated to frequencies of occurrence (Vitevitch and Luce 1998, 1999). Such effects of phonotactics are often gradient; at least, the empirical evidence has the form of observed facilitation of highly frequent combinations. On the other hand, there are also categorical effects, usually associated with application of phonotactic illegality, e.g., misperceptions of illegal phonotactic structures (Polivanov 1931, Hallé et al. 1998).

Gradience and categoricalness, but also formal knowledge and observed effects, have to be distinguished to be able to describe and explain phonotactic effects on perception. There is probabilistic knowledge, of gradient well-formedness, which formally expresses the degree to which a structure (e.g. phoneme combination) is phonotactically well-formed. There is also categorical knowledge of markedness, with certain marked structures being illegal. It is tempting to separate categorical effects as caused by categorical markedness and gradient effects by probabilistic knowledge, but such a separation has detrimental effects, as will be explained below.

The effects of phonotactic knowledge on speech perception are observed between acoustic input and lexical access. A perception grammar containing phonotactic knowledge should therefore be embedded in a model of word recognition as a preparser, mapping continuous speech input to phonological percepts that are subsequently mapped onto words by a lexical recognition grammar. Ruling out phonotactically illegal candidates is generally beneficial, hence effects of markedness can be categorical: phonotactically illegal phoneme strings can by definition not occur in actual words. After illegal options are discarded, multiple legal phonological percepts remain.

Instead of yielding one optimal percept, a grammar mapping acoustic input to phonological percepts must be capable of returning multiple candidates, with possible gradient differences in well-formedness.

2. How phonotactic knowledge can affect perception

2.1. Categorical markedness

This contribution describes the preparsing process with a grammar similar to Boersma's (1998, 1999) proposal, which employs the framework of Optimality Theory (OT; Prince and Smolensky 2002 [1993]). In the perception grammar, acoustic cue constraints govern the mapping between acoustic input and the phonemic level. Phonotactic cues have to be taken into account in perception as well to account for phonotactically driven misperceptions and perceptual biases (the theoretical possibility was suggested by Boersma 2007). Phonotactic constraints should be ranked among cue constraints that protect perfect matches and those that protect shoddy matches.

One form of gradience can also be described with categorical markedness: if the acoustic input is not close to any category, the perception grammar will still recognise it as a near category, but phonotactic cues can modulate this process towards the nearest phonotactically legal percept. Evidence that this modulation occurs is provided by e.g. Massaro and Cohen (1983). They used stimuli such as /ple/–/pre/, both legal, but also /tle/–/tre/, where the /l/ is not legal, or /sle/–/sre/, where the /r/ is not legal. Perception of the second phoneme depended on the F3 value at the onset of the glide. It was manipulated to be 2397, 2263, 2136, 2016, 1903, 1796 or 1695 Hz. The first is most like /l/, the last most like /r/.

To formalise this in the perception grammar, the acoustics have to be covered by cue constraints, such as a constraint *WARP([2136], 2397), militating against perceiving a stimulus with an F3 value of 2136 as belonging to the category of 2397 Hz. The latter category corresponds to the phoneme /l/; for ease of exposition the cue-to-category constraints will be expressed with reference to phonemes, hence *WARP([2136], /l/) expresses that the F3 value 2136 should not be categorised as /l/. There is also a constraint *WARP([2136], /r/) that forbids mapping input with a 2136 F3 value to /r/ and a general constraint against not categorising the input, CATEG. The perceptual input-output mapping of a stimulus with an F3 of 2136 Hz will proceed as shown in the tableau of Table *1* when the context is phonotactically neutral.

Table 1. Perception without phonotactics: ambiguous input

p[2136]e	CATEG	*WARP([2136], /l/)	*WARP([2136], /r/)
☞ /pre/			*
/p2136e/	*!		
/ple/		*!	

For ambiguous input, the ranking of the two constraints determines the best output, namely /r/; the constraint against categorising the 2136 Hz F3 as /l/ is evaluated first, as it is ranked highest and /r/ is perceived. When the input F3 is 2397, i.e. a perfect /l/, the tableau changes to the one in Table *2*, because the constraint against perceiving the input correctly as /l/, *WARP([2397], /l/), is ranked low and the constraint against perceiving the input as /r/, *WARP([2397], /r/), is high.

Table 2. Perception with perfect input. *WARP has been abbreviated to *W.

p[2397]e	CATEG	*W([2397], /r/)	*W([2136], /l/)	*W([2136], /r/)	*W([2397], /l/)
/pre/		*!			
/p2397e/	*!				
☞ /ple/					*

Massaro and Cohen (1983) found that the ambiguous stimuli were more likely to be perceived as the phoneme that was phonotactically legal, given the context. Hence, an ambiguous sound in the context s#e was more likely to be perceived as /l/ than as /r/. However, if s[2136]e is the input to the tableau in Table *1*, the outcome would still be /sre/. If the example of a sound with an F3 of 2136 is assumed to be subject to a phonotactic influence, a constraint against /sr/ is enough to model this effect, as shown in Table *3*. The ranking of */sr/ among the * WARP determines how strong the phonotactic effect is, in other words, the degree to which acoustic cues are ignored to fulfil the phonotactic constraint.

Table 3. Perception with phonotactics: ambiguous input. The option of not categorising and the constraint against it have been suppressed.

s[2136]e	*W([2397], /r/)	*/sr/	*W([2136], /l/)	*W([2136], /r/)	*W([2397], /l/)
/sre/		*!		*	
☞ /sle/			*		

The phonotactic constraint might not prevent illegal perceptions if the input is a perfect match with an illegal combination, as shown in Table *4*, now for the illegal combination /tl/, against which a constraint is also supposed to exist.

Table 4. Perception with phonotactics: perfect input.

t[2397]e	*W([2397], /r/)	*/tl/	*W([2136], /l/)	*W([2136], /r/)	*W([2397], /l/)
/tre/	*!				
☞ /tle/		*			*

Categorical effects can also be explained with a perception grammar containing phonotactic constraints. If the constraint */tl/ was ranked at the top of the grammar in Table *4*, the percept would never have been the illegal /tl/, however well the F3 matches the /l/.

Another type of gradience that OT grammars can capture is found in speech segmentation. The perception grammar should not only normalise acoustic input to phonemes, but also parse the continuous speech input into word-like chunks. The chunks might not be actual words; they are subject to lexical look-up. Under the assumption that there are (too) many (partial) matches between continuous speech input and lexical items at the phonological level, as well as the assumption that lexical look-up is costly, a preparser is useful as it directs lexical look-up to the optimal way to match input to words. If lexical look-up would be initiated at every phoneme in the speech input, many matches that later in the speech stream fail to continue matching will arise. Partial lexical look-up, as predicted by the Cohort model (Marslen-Wilson 1987), does occur, but phonotactic guiding is found for speech segmentation (Mattys et al. 2005, Lentz 2011: Ch. 3-4; cf. McQueen 1998).

Such guidance can partially be modelled by application of markedness constraints, e.g. by splitting illegal clusters with word boundaries, which is generally correct. E.g, McQueen (1998) found that a word such as /rɔk/ ‘rok’ *skirt* in Dutch is easier to spot when embedded after a consonant that cannot form a legal cluster with the word. E.g., in [fɪmrɔk], the word /rɔk/ was spotted more easily than in [fɪdrɔk]. In OT terms, the mapping [fɪmrɔk] to /fɪm.rɔk/ (the dot denotes a boundary) is better than other segmentations, because it does not violate */mr/. In the case of [fɪdrɔk], the segmentation /fɪ.drɔk/ is optimal, because it does not violate the Dutch constraint against voiced obstuents in the coda.

In normal language, the optimal candidate might also not always match lexical items. A phonotactically suboptimal parse might be lexically optimal. Boundaries might be missed, because they are not in phonotactically illegal contexts, or incorrectly inserted in marked but not fully illegal clusters. If the preparser is sound, its optimal candidate might be the most likely to be correct over a large number of similar parses, but not necessarily in every instance. If lexical look-up using the phonologically optimal percept fails, the preparser has to provide a new parse (Kager 2010). The more well-formed a segmentation is, the more likely it is to arise quickly; thus, this small adaptation to classic OT yields gradience in effects.

An illustration of the phonotactic grammar’s segmentation capacity is a description of the behaviour of participants in a Dutch word spotting study by Lentz (2011: Ch. 3, 5). Participants’ optimal segmentation for illegal clusters as [sr] was /s.r/. However, Dutch was a second language for one group of participants; their first language allowed all relevant clusters. Participants from this group were more likely to make the segmentation /.sr/ than native Dutch listeners, if no word started at the

/r/. This gradient difference can be modelled by assuming that if the optimal /s.r/ solution is rejected, the preparser comes up with /.sr/ as the next option. This segmentation violates possible contiguity constraint protecting them. The non-native listeners are assumed to have not yet ranked */sr/ as high as native listeners. Any constraint protecting /sr/, e.g. a constraint preferring consonant clusters in onsets, cannot be ranked lower than */sr/, because the listeners prefer to split [sr]. The native and non-native grammars, tuned to their preference for segmentations aligned with the constraint */sr/, do not yet describe the gradient difference between the groups. This kind of gradience, arising from categorical constraints, needs the OT grammar to yield gradient well-formedness. The next section discusses how this goal can be obtained.

2.2. Gradient differences for identical evaluations

Coetzee (2009) shows one way for an OT grammar to assign different levels of phonotactic well-formedness to different legal forms. He proposes that gradient well-formedness be derived from the harmonicity of the output, as defined by the markedness constraints. Two forms can be compared on their violations of markedness constraints: the highest markedness constraint that is violated by only one of the candidates decides which of the candidates is more well-formed. Another option to capture gradience in a phonotactic grammar based on categorical constraints is provided by Stochastic OT (Boersma 1997, Boersma and Hayes 2001). In Stochastic OT, the ranking of each constraint is drawn from a Gaussian distribution around a ranking value for every new evaluation. The variation in ranking per evaluation allows the evaluation to vary: if the ranking values of two constraints are near each other, their order is likely to be different in different evaluations. The closer conflicting constraints are, the higher the variability of the optimal candidate. Boersma and Hayes (2001) propose that the degree of well-formedness be derived from probability of occurrence and that categoricalness is simply a limit case.

Stochastic OT can solve the problem posed in the previous section. An output candidate has a probability of being optimal in a series of cases of the same input, but this does not directly translate to a well-formedness value in an individual case, unless the evaluation process is assumed to be repeated many times. The latter can be assumed to have happened in the segmentation experiments given above, as the results reported on these experiments are averaged over participants and items. If in the native listeners' grammar the */sr/ constraint is ranked higher than in the non-native listeners' grammar, the latter has the */sr/ closer to the competing contiguity constraint. In some evaluations, random noise on the ranking values will flip the order of the constraints and hence change the optimal percept, but this is more likely in the non-native grammar with the constraint */sr/ ranked closer to the contiguity constraint. The native grammar with the constraints ranked far apart predicts the native listeners to hardly ever come up with anything but /bes.ron/, while the non-native listeners with the constraints ranked closely sometimes evaluate the other option /be.sron/ as optimal (see Table *6*). Note that Coetzee's proposal, that does not feature stochastic rankings, cannot differ between the native and non-native well-formedness in this way, as the absolute ranking of the markedness constraint is the same in both grammars. However, if his proposal is adapted to derive well-formedness from the ranking value of violated constraints, the same difference in well-formedness of mappings containing /sr/ would be found between native and non-native listeners.

Table 6. Two ficticious Stochastic OT evaluations exemplifying segmentation with Slavic-influenced phonotactics. In this fictional example, the two constraints are ranked close, at 100 and 102. The actual ranking is decided anew for every evaluation by drawing a value from a Gaussian distribution around the ranking value of the constraints.

Evaluation 1

[besron]	102+2=104 */sr/	100−1=99 ONSET(CC)
/be.sron/	*!	
☞/bes.ron/		*
/besr.on/	*!	
/besron/	*!	

Evaluation 2

[besron]	100+1=101 ONSET(CC)	102−2=100 */sr/
☞/be.sron/		*
/bes.ron/	*!	
☞/besr.on/		*
☞/besron/		*

2.3. On the form of phonotactic knowledge

Only negative and categorical constraints have been used above. Probabilistic knowledge might be of a positive nature (the more a structure occurs, the easier it is to process). One might consider that even negative knowledge is in fact inverted positive knowledge, i.e. absence of positive knowledge causes markedness effects. However, Weber and Cutler (2006) report results for second language perception cannot be reduced to knowledge of legality, namely that listeners use knowledge of markedness from their first language, as well as from their second language, to segment speech. In German, /sl/ cannot occur within an onset, while this is legal in English. On the other hand, English forbids e.g. /ʃl/ in onsets, while this is legal in German. In both languages, both clusters cannot be codas. Participants in Weber and Cutler’s experiment were Germans who were highly proficient in English, as well as English native listeners that did not know German. They had to spot words in nonsense strings and the illegality of clusters provided a benefit to word recognition, because such clusters cue a boundary aligned with the word. E.g., in the case of the nonsense strings [darslidʒən] and [darʃlidʒən], containing the English word /lidʒən/, ‘legion’, the German learners used both English and German illegality knowledge and were able to benefit from it and spot the word with greater ease than in a baseline condition. English listeners did not consider /sl/ illegal, as it is not illegal in

English; therefore they do not benefit from facilitation provided by the German illegality of /sl/ that cues the segmentation /dars.lidʒən/, which does not contain the offensive combination and thus does not violate the constraint */sl/. The English listeners only spotted words with greater ease in the case of [darʃlidʒən], showing use of the knowledge that /ʃl/ is illegal in English. This suggests that the English knowledge that the German participants had acquired indeed resembles markedness contraints, namely */ʃl/, in other words that /ʃl/ is illegal. They would be predicted to have greater trouble, not less, if they would instead have acquired that /sl/ is more probable in English than in German. Speech segmentation thus provides an important indication for the existence of representations of phonotactic markedness that cannot be reduced to positive knowledge.

Additional evidence for the influence of phonotactic knowledge on speech segmentation can be found in Lentz (2011: Ch. 3), where pairs of illegal clusters were pitted against each other, e.g. /dl/ and /tl/, in which the first cannot be split as the result would be a voiced obstruent in coda position, which is illegal in Dutch, the experiment's language. Words (starting with /l/) are detected easier when preceded with a nonsense string ending on /t/ than if the string ended in /d/ (e.g. /lɑp/ 'lap' *rag* in [sytlɑp] is easier to spot than in [sydlɑp], /syt/ nor */syd/ being a word but the latter being phonotactically illegal).

In fact, markedness and probabilistic knowledge cannot be reduced to the same representation. Spanish-language learners of Dutch shows that misperceptions in Dutch, caused by the Spanish illegality of s+consonants (sC) clusters, does not exclude acquisition of the higher well-formedness in Dutch of some of these clusters (Lentz 2011: Ch. 2). Listeners can thus entertain conflicting probabilistic and markedness knowledge about the same structures. If the assumption that positive knowledge is necessary is licensed, the question remains whether it can be incorporated in the same grammar as the markedness.

If neither grammars based on categorical markedness constraints, nor grammars based on probabilistic knowledge seem to allow the prediction of gradient facilitatory effects as well as categorical effects, the assumption that both kinds of knowledge exist is licensed by the complexity of the influence of phonotactics on perception. A new question is then if both types of knowledge act within the same grammar, or if a markedness-based grammar explains categorical effects, separate from a probabilistic grammar that explains gradient effects. Nevertheless, a separation of markedness and probability together with respectively categorical and gradient effects is not desirable, as it makes it impossible to describe gradient effects caused by markedness.

Such descriptions are especially desirable in cross-language descriptions, as illustrated by cases in which a sound combination is avoided in one language depending on the phonological context, while it is categorically illegal in another language. Such a case is coda obstruent voicing in Dutch and German. Dutch generally does not allow coda obstruents to be voiced, but regressive voicing assimilation can optionally change a coda obstruent into a voiced one (Wetzels and Mascaró 2001, Zonneveld 2007). In other words, voiced obstruents are categorically illegal when assimilation does not apply, but when assimilation does apply, it has a probabilistic nature. German has completely categorical final devoicing, since it does not feature voice assimilation. This cross-linguistic observation shows that one and the same chunk of knowledge (the constraint against voiced obstruents in codas) can have gradient effects in one language and categorical effects in another. OT grammars containing a devoicing constraint capture devoicing with its phonological principle and phonetic properties, as the constraint is grounded in articulatory phonetics and confirmed by language typology; in addition, final devoicing is observed categorically in those contexts that are unaffected by contact with other phonemes in Dutch, indicating an interaction with other factors. A theory with a final devoicing constraint and constraints governing voice assimilation can explain the legality and thus non-zero probability of obstruent devoicing in both Dutch and German codas, as well as its categoricalness in German. The idea that linguistic regularities are observed in interaction with other regularities is the hallmark of constraint-based theories, in which constraints can be violated but only if needed to avoid the violation of more important constraints. To also incorporate the probabilistic nature of certain effects, such as voicing assimilation but also the phonotactic facilitation discussed above, classic OT grammars do not suffice, but adaptations of such as Stochastic OT do. This will be discussed below for the case of the gradient differences in ease of processing between forms of different probabilistic phonotactic well-formedness.

2.4. Positive knowledge within constraint-based grammars

The gradience associated with probabilistic phonotactics occurs between different mappings, i.e. between two or more mappings that do not share input. The output of each mapping is optimal, but it is reached more quickly when it is more well-formed: effects of this nature were reported by e.g. Vitevitch and Luce (1998, 1999) and Luce and Large (2001). These authors found that words containing frequent biphones are recognised faster than words containing less frequent biphones (when the confounding effects of lexical neighbourhood are taken into account). The mappings taking place in both cases do not violate faithfulness constraints, as both words were properly produced. A classical OT approach with markedness constraints expressing phonotactic knowledge thus does not suffice to explain such phonotactic effects on perception. Probabilistic effects might be better described with a non-markedness based knowledge representations. Vitevitch and Luce (1998, 1999) already proposed the existence of sublexical representational units for high-frequent phoneme combinations, next to word and phoneme representations. Such units serve to provide extra connectivity between input and word representations, aiding recognition of those words that contain the frequent phoneme combinations.

Gradient effects and probabilistic (positive) differences between different legal percepts are not captured in such classic OT grammars. Above, two ways in which OT grammars can be adapted to yield gradient effects as well have been discussed. However, these gradient effects were derived from markedness constraints, i.e. from negative knowledge and do not yet relate to positive probabilistic knowledge. They were also not for different input, but about the probability that a different mapping occurs for the same input.

Borrowing Vitevitch and Luce's idea that there are representations for highly well-formed phoneme combinations, constraints can be envisaged that refer to well-formed (highly probable) combinations, in a fashion similar to the contiguity constraints proposed by Adriaans and Kager (2010), who propose that the high probability of some phoneme sequences should be taken into account in speech segmentation. Contiguity constraints represent positive knowledge, but Adriaans and Kager's proposal interprets them as a ban on splitting well-formed structures. In this way, they function similarly to markedness constraints, in that they are assessed by violations, not by satisfactions. Hence, the standard evaluation of candidates with a ranked constraint set still applies.

As probabilistic phonotactic knowledge also affects word recognition in general, contiguity constraints should be interpreted more liberally as generally favouring perception of certain phonotactically well-formed structures. Contiguity constraints are then constraints against not recognising certain combinations. If such constraints force their biases on perception without mitigation, they would change the input to the most well-formed word without respect to the input. Consider as an illustration (borrowed from Vitevitch and Luce 1999) words with the high-frequent biphone /kæ/, such as 'cat', that are recognised more easily than words without high-frequent biphones such as 'fish'. Input of the form [fɪʃ] might then be more likely to be perceived as 'cat' than as 'fish', because the latter violates a constraint in favour of recognising /kæ/. Such abrupt changes in perception can be prevented by high-ranked faithfulness constraints for phonemes: constraints protecting the mapping [f]-/f/, [ɪ]-/ɪ/ and [ʃ]-/ʃ/ stop the recognition of [fɪʃ] as 'cat', when ranked above the /kæ/ constraint. Alternatively, one can assume contiguity constraints to be faithfulness constraints themselves, referring to input-output mappings. If the constraint /kæ/ is interpreted as "do not perceive anything without /kæ/ if the input contains [kæ]", wild misperceptions are also avoided, without the necessary assumptions that other constraints outrank them.

Contiguity constraints such as /kæ/ should make the mapping [kæt]–/kæt/ more optimal than mappings with words containing less probable biphones, e.g. /fɪʃ/ 'fish'. In the gradient well-formedness calculation of Coetzee (2009), both are assessed on the same set of markedness constraints, so if [fɪʃ]–/fɪʃ/ violates more important markedness constraints than [kæt]–/kæt/ the gradient difference is covered. In the simple reading of /kæ/ as "perceive /kæ/", the perceptual mapping [fɪʃ]-/fɪʃ/ violates /kæ/. Under the reading that the contiguity constraint /kæ/ protects the biphone when present in the input, Coetzee's proposal to derive gradient well-formedness differences would not work, because [fɪʃ]-/fɪʃ/ then also fulfils this constraint (vacuously).

If Stochastic OT is used to generate gradient effects, the form of the contiguity constraint can be both. If a general constraint to perceive well-formed structures is used, with a ranking value derived from the probability of the structure, the more probable a structure is, the higher ranked the constraint requiring its perception. This makes percepts containing these structures more likely to be optimal, as they do not violate the high-ranked contiguity constraint. If probabilistic well-formedness is seen as captured with faithfulness constraints, in an evaluation in which the drawn rankings of the faithfulness constraints protecting the individual phonemes are low, the biphone faithfulness constraint is still probably ranked high, also explaining ease of processing compared to cases in which the input does not containing highly probable structures. In the latter case, constraints against categorisation of input can be ranked higher than phoneme faithfulness constraints;

the ranking of e.g. biphone constraints are not likely to come to the rescue if they are not ranked high.

4. Conclusion

An OT grammar very similar (if not equal) to Boersma's (1999) proposal, either as Stochastic OT applied repeatedly or as classic OT with additional calculation of gradient well-formedness using Coetzee's (2009) proposal, can render categorical effects such as misperceptions, gradient effects associated with illegality or markedness as described above for speech segmentation and the gradient facilitation in processing associated with probabilistic phonotactics.

If a constraint-based grammar of the kind described above performs the evaluation of output candidates against the set of constraints representing phonotactic knowledge, speech segmentation, perceptual illusions and phonotactic influences on the perception of sequences of ambiguous sound that occur before lexical recognition can be described; the perceptual output is a ranked set of candidates on a level before lexical access. The highest-ranked candidate is not optimal in the classic sense as being the only grammatical option, but it is optimal in a more everyday use of that term: it is the solution that is most probably the correct one. As it is not guaranteed to be the only solution, less well-formed candidates can also be submitted to lexical look-up. In this process, probabilistic knowledge of frequent sound combinations has gradient effects. Phonotactic knowledge aids the speech recognition process at the prelexical level by mapping raw acoustic input to a string of pseudowords, normalised by phonetic and phonotactic constraints in the perceptual grammar. The preparser restricts the hypothesis space for word recognition and thus allows lexical material to connect more efficiently to the input, bypassing unnecessary computations such as the lexical look-up of a pseudoword with an illegal sound combination.

References

Adriaans, Frans & René Kager. 2010. Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language* 62. 311–331.

Auer, Edward T. & Paul A. Luce, 2005. Probabilistic phonotactics in spoken word recognition. In David B. Pisoni and Robert E. Remez (eds), *The handbook of speech perception*, 610–630. Malden, MA: Blackwell.

Berent, Iris, Donca Steriade, Tracy Lennertz & Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, 104(3). 591–630.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. In *Proceedings of the Institute of Phonetic Science, University of Amsterdam* 21. 43–58.

Boersma, Paul. 1998. *Functional phonology. Formalizing the interactions between articulatory and perceptual drives* 11. The Hague: Holland Academic Graphics.

Boersma, Paul. 1999. *On the need for a separate perception grammar.* Online. roa.rutgers.edu. (ROA-358)

Boersma, Paul 2007. Some listener-oriented accounts of *h*-aspiré in French. *Lingua* 117(12). 1989–2054.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32(1). 45–86.

Brown, Roger, & Donald Hildum. 1956. Expectancy and the perception of syllables. *Language* 32. 411–419.

Chomsky, Noam & Morris Halle. 1965. Some controversial questions in phonological theory. *Journal of Linguistics* 1. 97–138.

Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84(2). 218–257.

Coetzee, Andries W. 2009. Grammar is both categorical and gradient. In Steven Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, 9–42. London: Equinox.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. Epenthetic Vowels in Japanese: A Perceptual Illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25(6). 1568–1578.

Hallé, Pierre A., Juan Segui, Uli Frauenfelder & Christine Meunier. 1998. Processing of illegal consonant clusters: a case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24(2). 592–608.

Kager, René. 2010. An Optimality-Theorety perspective on speech segmentation. *Journal of the Phonetic Society of Japan* 14(1). 24–34.

Lentz, Tom. 2011. *Phonotactic illegality and probability in speech perception: Evidence from second language listeners.* Utrecht University dissertation. Utrecht: LOT.

Luce, Paul A. & Nathan R. Large. 2001. Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes* 16(5). 565–581.

Marslen-Wilson, William. 1987. Functional parallelism in spoken word recognition. *Cognition* 25. 71–102.

Massaro, Dominic W., & Michael M. Cohen. 1983. Phonological context in speech perception. *Perception and Psychophysics* 34(4). 338–348.

Mattys, Sven L., Laurence White & James F. Melhorn. 2005. Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General* 134(4). 477–500.

McQueen, James M. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39. 21–46.

Moreton, Elliott. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84. 55–71.

Norris, Dennis, James M. McQueen, Anne Cutler & Sally Butterfield. 1997. The Possible-Word Constraint in the segmentation of continuous speech. *Cognitive Psychology* 34. 191–243.

Polivanov, Evgenij. 1931. La perception des sons d'une langue étrangère [Perception of sounds of a foreign language]. In *Travaux du cercle linguistique de Prague: Vol. 4. réunion phonologique internationale tenue a Prague (18–*

21/xii 1930), 79–96. Prague: Jednota Československých Matematiků a Fysiků.

Prince, Alan & Paul Smolensky. 2004 [1993]. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell.

Vitevitch, Michael S. & Paul A. Luce. 1998. When words compete: Levels of processing in perception of spoken words. *Psychological Science* 9(4). 325–329.

Vitevitch, Michael S. & Paul A. Luce. 1999. Probabilistic phonotactics and neighbourhood activation in spoken word recognition. *Journal of Memory and Language* 40. 374–408.

Weber, Andrea & Anne Cutler. 2006. First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America* 119(1). 597–607.

Wetzels, W. Leo & Joan Mascaró. 2001. The typology of voicing and devoicing. *Language* 77(2). 207–244.

Zonneveld, Wim. 2007. Issues in Dutch devoicing: Positional faithfulness, positional markedness, and local conjunction. In Jeroen van de Weijer & Erik Jan van der Torre (eds.), *Voicing in Dutch: (De)voicing - phonology, phonetics, and psycholinguistics* 1–40. Amsterdam / Philadelphia: John Benjamins.